

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2005, 23:46:57 ; Search time 8541.75 Seconds
(without alignments)
11177.792 Million cell updates/sec

Title: US-10-069-056-4
Perfect score: 2019
Sequence: 1 agcctggaatgcttactc.....agccgtgaactggactaa 2019

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_cm.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2019	100.0	2019	6	AX137739 Sequence
2	2017.4	99.9	2019	6	AX137736 Sequence
3	2017.4	99.9	5081	14	PAMVM2
4	2017.4	99.9	5149	14	MVMPCG
5	2015.8	99.8	2019	6	AX137743 Sequence
6	2015.8	99.8	2019	6	AX137747 Sequence
7	2015.8	99.8	2019	6	AX137751 Sequence
8	1902.2	94.4	5085	14	MVMICG
9	1902.2	94.2	4764	14	MU34253
10	1902.2	94.2	5087	14	PAMVMI
11	1899	94.1	5144	14	MPU12469
12	1873	92.8	4764	14	MU34254
13	1862.2	92.2	4773	14	HU34255
14	1855.8	91.9	4761	14	MU34256
15	1771	87.7	5121	6	CQ786765
16	1662.2	82.3	4904	14	AF321230
17	1659	82.2	4795	14	AF321230
18	1657.4	82.1	4927	14	KRU79033
19	1656.6	82.1	5135	14	PVRSEQ

20	1655.8	82.0	4816	14	AF332883	AF332883 Rat minut
21	1652.6	81.9	4813	14	AF332882	AF332882 Rat minut
22	1644.6	81.5	5176	14	PARH1	X01457 Parvovirus
23	1644.4	81.4	3995	14	AF036711	AF036711 Kilham ra
24	1275	63.2	4936	14	AF036710	AF036710 Rat parvo
25	1051.6	52.1	2007	14	AB000062	AB000062 Feline pa
26	1050	52.0	2007	14	AB000048	AB000048 Feline pa
27	1050	52.0	2007	14	AB000057	AB000057 Feline pa
28	1050	52.0	5124	14	PVPVC	M38246 Feline panl
29	1046.8	51.8	2007	14	AB000049	AB000049 Feline pa
30	1046.8	51.8	2007	14	AB000058	AB000058 Feline pa
31	1046.8	51.8	2007	14	AB000060	AB000060 Feline pa
32	1046.8	51.8	2007	14	AB000069	AB000069 Feline pa
33	1045.2	51.8	2007	14	AB000053	AB000053 Feline pa
34	1045.2	51.8	2007	14	AB000063	AB000063 Feline pa
35	1045.2	51.8	2007	14	AB000065	AB000065 Feline pa
36	1045.2	51.8	2007	14	AB000067	AB000067 Feline pa
37	1045.2	51.8	5075	14	PVCY1A	D26079 Canine parv
38	1043.6	51.7	2007	14	AB000051	AB000051 Feline pa
39	1043.6	51.7	4983	14	PAFPVNS1	X55115 Feline panl
40	1043.6	51.7	5124	14	PVCPVC	M38245 Canine parv
41	1042	51.6	2007	14	AB000055	AB000055 Feline pa
42	1042	51.6	5323	14	PVCCPN	M19296 Canine parv
43	1040.4	51.5	5049	6	AR043629	AR043629 Sequence
44	1040.4	51.5	5049	6	AR043630	AR043630 Sequence
45	1040.4	51.5	5094	14	PVMRFD	D00765 Mink enteri

ALIGNMENTS

RESULT 1	AX137739	Sequence 4 from Patent EP1077260.	2019 bp	DNA	linear	PAT 30-MAY-2001
AX137739	LOCUS					
AX137739	DEFINITION					
AX137739	ACCESSION					
AX137739.1	VERSION	GI:14273913				
MIce minute virus	KEYWORDS					
MIce minute virus	SOURCE					
Viruses; sadna viruses; Parvoviridae; Parvovirinae; Parvovirus.	ORGANISM					
Nuesch,J. and Rommelaere,J.	REFERENCE					
Parvovirus ns1 variants	AUTHORS					
Patent: EP 1077260-A 4 21-FEB-2001;	TITLE					
Deutsches Krebsforschungszentrum' Stiftung des Oeffentlichen Rechts	JOURNAL					
(DE)						
Location/Qualifiers	FEATURES					
1..2019	source					
/organism="Mice minute virus"						
/mol_type="unassigned DNA"						
/db_xref="taxon:10794"						
/note="Parvovirus NS1 variant"						
1..2019						
/notes="unnamed protein product"						
/codon_start=1						
/protein_id="CAC39991.1"						
/db_xref="GI:14273914"						
/translation="MAGNAYSDEVLTGATNWLKEKSNQEVFSVFVKENVOLNGDKIGW						
NSYKELQDELSLQRGAEATWQSDQEDWETTVDMETTKQVFI FDSLVKKLCEVL						
NTKNIFPGDNVNFVQHEWQDQGHCHVLIGGKDFQQAQGWRRQLNVVWSRLVTA						
CNVQTPAERI KLR EIAEDNFWLLTYKHQTKDYTKCVLFGNMAYYFLTKKGIS						
TSPPRDGGYFLSDSGWTKNFKEGEHLVSKLYTDDMRPETVETVTTVAQETKRGRI						
OTKEVAI KTLLEKLVHRTVSPEDMMWQPSYIEMWAQPGGELNKNLETCLETLTL						
ARTKAPDLLEKAEFTSLTNESLPDTRTCRIFAFGWYVKVCHACCVLNQGGKR						
NTVLFHGASTGKSIIFQAIAQAGNVGNCYNAANVPNDCTNKNLIWEEAGNFGQ						
QYNQFKACSGGTIRIDQKGSQKLEPTFVIMTINENITVVRIGCEERPEHQPIRD						
RLMLHLTLTDFGLVDKNWPMICAWLVKNGYQSTWASVCARKGKVPDWSENNAE						
PKVPTPINLLGSARSPTTPKSTPLSQNTALTPLASDLEDLALPESTNTPTVAGTAE						
TQNTGTGAGSKAQDQGLSPTWSEIEEDLRACFGAEPLKKDSEPLND"						

ORIGIN

Query Match 100.0%; Score 2019; DB 6; Length 2019;

Best Local Similarity 100.0%; Pred. No. 0; Matches 2019; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	ATGGCTGGAATGCTTACTCTGATGAAGTTTTCGGAGCAACCAACTGGTTAAAGCAAAA	60						
Db	1	ATGGCTGGAATGCTTACTCTGATGAAGTTTTCGGAGCAACCAACTGGTTAAAGCAAAA	60						
Qy	61	AGTAACCAAGGAGTGTTCTCATTTGGTTTTTAAATAATGAAATGTTCAACTGAATGAAAA	120						
Db	61	AGTAACCAAGGAGTGTTCTCATTTGGTTTTTAAATAATGAAATGTTCAACTGAATGAAAA	120						
Qy	121	GATATCGGATGAATAGTTTACAAAAAGAGCTCGAGGAGGACGAGCTGAATCTTTTACAA	180						
Db	121	GATATCGGATGAATAGTTTACAAAAAGAGCTCGAGGAGGAGGAGCTGAATCTTTTACAA	180						
Qy	181	CGAGGCGGAAACTACTCTGGGACCAAGCGAGGACATGGAATGGGAACCAACAGTGGAT	240						
Db	181	CGAGGCGGAAACTACTCTGGGACCAAGCGAGGACATGGAATGGGAACCAACAGTGGAT	240						
Qy	241	GAATGACCAAAAGCAAGTATTTCAITTTTGGTTTAAAAAATGTTTATTGAA	300						
Db	241	GAATGACCAAAAGCAAGTATTTCAITTTTGGTTTAAAAAATGTTTATTGAA	300						
Qy	301	GTGCTTAAACAAGAAATATATTTCTCTGTGTATGTTAAATGTTTGTGCAACATGAATGG	360						
Db	301	GTGCTTAAACAAGAAATATATTTCTCTGTGTATGTTAAATGTTTGTGCAACATGAATGG	360						
Qy	361	GGAAAAACCAAGCTGGCACTGCCATGTACTAAATGGAGGAAAGGACTTTAGTCAAGCT	420						
Db	361	GGAAAAACCAAGCTGGCACTGCCATGTACTAAATGGAGGAAAGGACTTTAGTCAAGCT	420						
Qy	421	CAAGGGAATGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTGTAAACAGCC	480						
Db	421	CAAGGGAATGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTGTAAACAGCC	480						
Qy	481	TGTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGACAAT	540						
Db	481	TGTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGACAAT	540						
Qy	541	GAGTGGGTTACTCTACTTTATAGCATAGCAAAACCAAAAAGACTATACCAAGTGT	600						
Db	541	GAGTGGGTTACTCTACTTTATAGCATAGCAAAACCAAAAAGACTATACCAAGTGT	600						
Qy	601	GTTCCTTTTGGAAACATGATGTGTTACTATTTTAACTTAAAAAGAAATTAAGCACTAGT	660						
Db	601	GTTCCTTTTGGAAACATGATGTGTTACTATTTTAACTTAAAAAGAAATTAAGCACTAGT	660						
Qy	661	CCACCAAGAGACGGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA	720						
Db	661	CCACCAAGAGACGGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA	720						
Qy	721	AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACATGATGACATCGCGCCAGAAACG	780						
Db	721	AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACATGATGACATCGCGCCAGAAACG	780						
Qy	781	GTTGAAACCAAGTAACTACCTGCGAGGAAACTAAGCGCGGAGAAATTCAACTTAAAAA	840						
Db	781	GTTGAAACCAAGTAACTACCTGCGAGGAAACTAAGCGCGGAGAAATTCAACTTAAAAA	840						
Qy	841	GAAAGTTGCTATTAAAACTACACTTAAAGAGCTGGTGCAATAAAGAGTAACTCCACAGAG	900						
Db	841	GAAAGTTGCTATTAAAACTACACTTAAAGAGCTGGTGCAATAAAGAGTAACTCCACAGAG	900						
Qy	901	GACTGGATGATGATGAGCGGACAGTACATTTGAATGATGGCTCAACAGGTGGAGAA	960						
Db	901	GACTGGATGATGATGAGCGGACAGTACATTTGAATGATGGCTCAACAGGTGGAGAA	960						
Qy	961	AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAACCAAAAACAGCA	1020						
Db	961	AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAACCAAAAACAGCA	1020						
Qy	1021	TTTGACTTAATTTTGAAGAAAGCTGAAACCAAGCAAACTTAAACCACTTTTCACTGCCTGAC	1080						

KEYWORDS	Mice minute virus	Db	481	TGTAATGTGCAACTAACACAGCTGAAAGAAATTAATACTAAGAGAAATAGCAGAAGCAAT	540
SOURCE	Mice minute virus	Qy	541	GAGTGGGTACTCTACTTACTTATATAGCATATAGCAAAACCAAAAGACTATACCAAGTGT	600
ORGANISM	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.	Db	541	GAGTGGGTACTCTACTTACTTATATAGCATATAGCAAAACCAAAAGACTATACCAAGTGT	600
REFERENCE	1 Nueesch,J. and Rommelaere, J.	Qy	601	GTTCCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAAAATAAGCACTAGT	660
AUTHORS	Parvovirus nel variants	Db	601	GTTCCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAAAATAAGCACTAGT	660
TITLE	Patent: EP 1077260-A 1 21-FEB-2001;	Qy	661	CCACCAAGAGACGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA	720
JOURNAL	Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts (DE)	Db	661	CCACCAAGAGACGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA	720
FEATURES	Location/Qualifiers	Qy	721	AAAGAGGCGAGGCGCCATCTAGTGAGCAAACTATACACTGATACATGCGGCCAGAAACG	780
source	1. .2019	Db	721	AAAGAGGCGAGGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG	780
CDS	/db_xref="taxon:10794"	Qy	781	GTTGAAACCAACAGTAAACCACTGCGCAGGAACTTAAGCGGCGAGAAATTCAAAATTA	840
	/note="Wildtype Parvovirus NS1"	Db	781	GTTGAAACCAACAGTAAACCACTGCGCAGGAACTTAAGCGGCGAGAAATTCAAAATTA	840
	/note="unnamed protein product"	Qy	841	GAAAGTTGCTATTAAAACTACATTTAAAGAGCTGGTGCATAAAAAGAGTAACCTCACCAGAG	900
	/protein_id="CAC39989.1"	Db	841	GAAAGTTGCTATTAAAACTACATTTAAAGAGCTGGTGCATAAAAAGAGTAACCTCACCAGAG	900
	/db_xref="GI:14273910"	Qy	901	GACTGATGATGATGAGCAGACAGTATACATTTGAAATGATGCTCAACAGGTGGAGAA	960
	/translation="MAGNAYSDVLGATNWLKESNQVFSFVFKNNVOLNGKIDIGW	Db	901	GACTGATGATGATGAGCAGACAGTATACATTTGAAATGATGCTCAACAGGTGGAGAA	960
	NSYKLEQDELKSLQGEATWQSDMEWETVDEMTKQVFIIDSLVKCLPEVL	Qy	961	AACTGCTGAAAATAACGCTAGAGATTTGTACACTACTCTGCCAGAACCAAAACAGCA	1020
	NTKNIIPGDVNVFVQHEWKGDCQCHVLIGKDFSOAGCKWRROLNVYWSRLVTA	Db	961	AACTGCTGAAAATAACGCTAGAGATTTGTACACTACTCTGCCAGAACCAAAACAGCA	1020
	CNVQLTPAERIKLREAEADNEWTLTYKHQTKDYTKCVLPGNMIAYFLTKKIS	Qy	1021	TTTGACTTAATTTTAAAAAGCTGAAACCCAGCAAACTAACCACTTTTCACTGCCTGAC	1080
	TSPPRGVYLSDSQSKTNFLKEGERHLVSKUYTDMPETVETITVTAQETKGR	Db	1021	TTTGACTTAATTTTAAAAAGCTGAAACCCAGCAAACTAACCACTTTTCACTGCCTGAC	1080
	QTKPEVSIKTTLELKHVRVTSPEDEMMQPDYSIEMAOQPGGENLLKNTLEICTL	Qy	1081	ACAAGAACCTGCGAGAAATTTTGGCTTTTCATGCTGGAACTATGTTTAAAGTTTGCATGCT	1140
	ARTKPAFLILEAETS KLTNFSLPDTRICRIFAFHGNVYKVCIAICVILNRQGGK	Db	1081	ACAAGAACCTGCGAGAAATTTTGGCTTTTCATGCTGGAACTATGTTTAAAGTTTGCATGCT	1140
	NWLFPGPASTGKSIATAIAQAVGNVCYNAANVFPNDCTNKNLIWVEEAGNGQ	Qy	1141	ATTGCTGTGTTTAAACAGACAAAGGAGGCAAAAGAAATACTGTTTATTTATTCAGACCA	1200
	OVNQFAISGOTIRIDOKGSKQIEPTVINTNENITVVRIGCEERPEHTQPIRD	Db	1141	ATTGCTGTGTTTAAACAGACAAAGGAGGCAAAAGAAATACTGTTTATTTATTCAGACCA	1200
	RLNLIHLTHLPDGLVDKNPIMICAMLVKNGYQSTWASCYAKWKGYPDPNSWAE	Qy	1201	GCCAGCACAGGCAAACTCTATTATTGCAAGCCATAGCACAAAGCAGTTGGCAATGTTGGT	1260
	KVPTPINLIGSASFPTPKSTPLSNVATPLASDLALPEPSTNTPVAGTAE	Db	1201	GCCAGCACAGGCAAACTCTATTATTGCAAGCCATAGCACAAAGCAGTTGGCAATGTTGGT	1260
	TONTGAGSKACQDGLSPTWSEIBEDLRACFGABPKKDFSEPLND	Qy	1261	TGCTATAATGCGACCAATGTAAACTTTCCATTTTAAATGACTGTACCAACAAGAACTTGATT	1320
ORIGIN		Db	1261	TGCTATAATGCGACCAATGTAAACTTTCCATTTTAAATGACTGTACCAACAAGAACTTGATT	1320
Query Match	99.9%; Score 2017.4; DB 6; Length 2019;	Qy	1321	TGGGTAGAAAGAGCTGGTAACTTTTGACAGCAAGTAAACCCAGTTTAAAGCACTTGCTCT	1380
Best Local Similarity	100.0%; Pred. No. 0;	Db	1321	TGGGTAGAAAGAGCTGGTAACTTTTGACAGCAAGTAAACCCAGTTTAAAGCACTTGCTCT	1380
Matches 2018; Conservative	0; Mismatches 1; Indels 0; Gaps 0;	Qy	1381	GGTCAAATATTTCGCACTTGATCAAAAAGGAAAGGAGCAAAACAGATTGAAACCAACCA	1440
Qy	1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGACCAACCACTGTTTAAAGGAAAA	60	1381	GGTCAAATATTTCGCACTTGATCAAAAAGGAAAGGAGCAAAACAGATTGAAACCAACCA	1440
Db	1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGACCAACCACTGTTTAAAGGAAAA	60	1441	GTCAATCATGACCAAAATGAGAACATTACAGTGGTTCAGAAATAGGCTGCGAAGAAAGACCA	1500
Qy	61 AGTAACGAGAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGAAAA	120	1501	GATCAGCTCAACCAATCAGAGACAGAAATGCTTAAATTCATCTACACATACCTTGCTCT	1560
Db	61 AGTAACGAGAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGAAAA	120	1501	GATCAGCTCAACCAATCAGAGACAGAAATGCTTAAATTCATCTACACATACCTTGCTCT	1560
Qy	121 GATATCGGATGGAATAGTTTACAAAAGAGCTGCGAGGACGAGCTGAAATCTTTACAA	180	1561	GGTGACCTTTGGTTGGTTCGCAAAAATGAATGCGCCATGATTTGCTGCTGTTGGTAAAG	1620
Db	121 GATATCGGATGGAATAGTTTACAAAAGAGCTGCGAGGACGAGCTGAAATCTTTACAA	180	1561	GGTGACCTTTGGTTGGTTCGCAAAAATGAATGCGCCATGATTTGCTGCTGTTGGTAAAG	1620
Qy	181 CGAGGAGCGAACTACTTGGACCAAGCAGACATGGAATGGGAAACCAAGTGGAT	240			
Db	181 CGAGGAGCGAACTACTTGGACCAAGCAGACATGGAATGGGAAACCAAGTGGAT	240			
Qy	241 GAAATGACCAAAAAGCAAGTATTCAATTTTGAATCTTTGGTTAAAAAATGTTTATTTGAA	300			
Db	241 GAAATGACCAAAAAGCAAGTATTCAATTTTGAATCTTTGGTTAAAAAATGTTTATTTGAA	300			
Qy	301 GTGCTTAAACCAAAAGAAATATTTCTCTGGTGATGTTAAATTTGTTGCAACATGAATGG	360			
Db	301 GTGCTTAAACCAAAAGAAATATTTCTCTGGTGATGTTAAATTTGTTGCAACATGAATGG	360			
Qy	361 GGAAGACCAAGCTGCGACTGCGATGTAATTTGGAGGAAAGGACTTTTGTCAAGCT	420			
Db	361 GGAAGACCAAGCTGCGACTGCGATGTAATTTGGAGGAAAGGACTTTTGTCAAGCT	420			
Qy	421 CAAGGGAATGTTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGTTGGTAAACAGCC	480			
Db	421 CAAGGGAATGTTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGTTGGTAAACAGCC	480			
Qy	481 TGTAATGTGCAACTAACACAGCTGAAAGAAATTAATACTAAGAGAAATAGCAGAAGCAAT	540			

Db 561 GTGCTTAAACAAGAAATATTTCTCGGTGATGTTAAATGGTTTGTGCAACATGAATGG 620
Qy 361 GGAAGAAGCAAGGCTGGCACTGTCATGTACTAATTTGGAGGAAGGACTTTAGTCAAGCT 420
Db 621 GGAAGAAGCAAGGCTGGCACTGTCATGTACTAATTTGGAGGAAGGACTTTAGTCAAGCT 680
Qy 421 CAAGGGAATGTGGAGGAAGCAATTAATTTTACCTGGAGCAGATGGTGGTAAACAGCC 480
Db 681 CAAGGGAATGTGGAGGAAGCAATTAATTTTACCTGGAGCAGATGGTGGTAAACAGCC 740
Qy 481 TGTAAATGTGCAACTAACACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGACAAT 540
Db 741 TGTAAATGTGCAACTAACACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGACAAT 800
Qy 541 GAGTGGGTACTCTACTTACTTATAAGCATAAGCAACCAAAAGAGACTATACCAAGTGT 600
Db 801 GAGTGGGTACTCTACTTACTTATAAGCATAAGCACAACCAAAAGAGACTATACCAAGTGT 860
Qy 601 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAGAAATTAAGCAGTGT 660
Db 861 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAGAAATTAAGCAGTGT 920
Qy 661 CCACCAAGAGACGGAGCTATTTCTTAGCAGTGACTCTGGCTGGAAGCACTTAACTTTTA 720
Db 921 CCACCAAGAGACGGAGCTATTTCTTAGCAGTGACTCTGGCTGGAAGCACTTAACTTTTA 980
Qy 721 AAAGAAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCGAGAAACG 780
Db 981 AAAGAAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCGAGAAACG 1040
Qy 781 GTTGAACACACAGTAACCACTGGCAGGAACTAAGCGCGGAGAAATTAACCAAAATAA 840
Db 1041 GTTGAACACACAGTAACCACTGGCAGGAACTAAGCGCGGAGAAATTAACCAAAATAA 1100
Qy 841 GAAGTGTCTATTAAACTACACTTAAGAGCTGGTGATTAAGAGCTTAACCTCACCAGAG 900
Db 1101 GAAGTGTCTATTAAACTACACTTAAGAGCTGGTGATTAAGAGCTTAACCTCACCAGAG 1160
Qy 901 GACTGGATGATGTCAGCAGCAGACAGTTACATTAATGAATGCTCAACAGAGTGGAGAA 960
Db 1161 GACTGGATGATGTCAGCAGCAGACAGTTACATTAATGAATGCTCAACAGAGTGGAGAA 1220
Qy 961 AACCTGCTGAAAATACGCTAGAGATTTGTACACTAACTAGCCAGAACCAAAACAGCA 1020
Db 1221 AACCTGCTGAAAATACGCTAGAGATTTGTACACTAACTAGCCAGAACCAAAACAGCA 1280
Qy 1021 TTTGACTTAATTTAGAAAAGCTGAACACGCAAACTAACCACTTTTCACTGCCTGAC 1080
Db 1281 TTTGACTTAATTTAGAAAAGCTGAACACGCAAACTAACCACTTTTCACTGCCTGAC 1340
Qy 1081 ACAAGAACCTGCAGAAATTTTGTCTTTTCATGGCTGGAATCTATGTTAAAGTTTGGCATGCT 1140
Db 1341 ACAAGAACCTGCAGAAATTTTGTCTTTTCATGGCTGGAATCTATGTTAAAGTTTGGCATGCT 1400
Qy 1141 ATTTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATCTGTTTATTTATTCATGACCA 1200
Db 1401 ATTTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATCTGTTTATTTATTCATGACCA 1460
Qy 1201 GCAGACAGGCAAACTCTATTTATGCACAAGCCATAGCAAGCAGTGGCAATGTTGGT 1260
Db 1461 GCAGACAGGCAAACTCTATTTATGCACAAGCCATAGCAAGCAGTGGCAATGTTGGT 1520
Qy 1261 TGCTATAATCAGCAATGTAACTTTTCCATTTAATGACTGTACCAACAGAACTTGATT 1320
Db 1521 TGCTATAATCAGCAATGTAACTTTTCCATTTAATGACTGTACCAACAGAACTTGATT 1580
Qy 1321 TGGGTAGAAGAGCTGGTAACTTTTGGACAGCAAGTAAACAGATTTAAAGCACTTTGCTCT 1380
Db 1581 TGGGTAGAAGAGCTGGTAACTTTTGGACAGCAAGTAAACAGATTTAAAGCACTTTGCTCT 1640
Qy 1381 GGTCAAACATTTGCTATGATCAAAAAGGAAAGGAGCAAAACAGATTTGAACCAACACCA 1440

Db 1641 GGTCAAACATTTGCTATGATCAAAAAGGAAAGGCAAGCAAGATTTGAACCAACACCA 1700
Qy 1441 GTCATCATGACCAAAATGAGAACATTTACAGTGGTCAAGATAGCTCGAAGAAAGACCA 1500
Db 1701 GTCATCATGACCAAAATGAGAACATTTACAGTGGTCAAGATAGCTCGAAGAAAGACCA 1760
Qy 1501 GAACACACTCAACCAATCAGAGACAGANTGCTTTAACTTCATCTAACACATACCTTGCT 1560
Db 1761 GAACACACTCAACCAATCAGAGACAGANTGCTTTAACTTCATCTAACACATACCTTGCT 1820
Qy 1561 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGTCTTGGTTGGTAAAG 1620
Db 1821 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGTCTTGGTTGGTAAAG 1880
Qy 1621 AATGGTTTCAATCTACATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTGATTGG 1680
Db 1881 AATGGTTTCAATCTACATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTGATTGG 1940
Qy 1681 TCAGAAAACTGGGGGAGCCAAAGGTGCCAACTTCTATTAATTTACTAGTTTGGGCAAGC 1740
Db 1941 TCAGAAAACTGGGGGAGCCAAAGGTGCCAACTTCTATTAATTTACTAGTTTGGGCAAGC 2000
Qy 1741 TCACCAATTCACGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTAACTTCCACTT 1800
Db 2001 TCACCAATTCACGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTAACTTCCACTT 2060
Qy 1801 GCATCGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAAACTATCTCTGTTGCG 1860
Db 2061 GCATCGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAAACTATCTCTGTTGCG 2120
Qy 1861 GGCATCTGAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920
Db 2121 GGCATCTGAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 2180
Qy 1921 CTGAGCCCACTTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTCGGTGGGAACCG 1980
Db 2181 CTGAGCCCACTTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTCGGTGGGAACCG 2240
Qy 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2019
Db 2241 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2279

RESULT 4

LOCUS MWMPG 5149 bp ss-DNA linear VRL 22-MAY-1995
DEFINITION Minute virus of mice, complete genome.
ACCESSION J02275 M12520 M12521 M14704
VERSION J02275.1 GI:332293
KEYWORDS alternative splicing; capsid protein; complete genome; nonstructural protein.
SOURCE Mice minute virus
ORGANISM Mice minute virus
REFERENCE 1 (bases 1 to 5149)
AUTHORS Astell,C.R., Thomson,M., Merchinsky,M. and Ward,D.C.
TITLE The complete DNA sequence of minute virus of mice, an autonomous parvovirus
JOURNAL Nucleic Acids Res. 11 (4), 999-1018 (1983)
MEDLINE 83143341
PUBMED 6298737
REFERENCE 2 (bases 1 to 5149)
AUTHORS Astell,C.R., Gardiner,E.M. and Tattersall,P.
TITLE DNA sequence of the lymphotropic variant of minute virus of mice, MVM(i), and comparison with the DNA sequence of the fibrotropic prototype strain
J. Virol. 57 (2), 656-669 (1986)
MEDLINE 86115415
PUBMED 3502703
REFERENCE 3 (sites)
AUTHORS Morgan,W.R. and Ward,D.C.
TITLE Three splicing patterns are used to excise the small intron common to all minute virus of mice RNAs

JOURNAL
MEDLINE
PUBMED
COMMENT

J. Virol. 60 (3), 1170-1174 (1986)

87061199

3783817

Original source text: Minute virus of mice (strain MVM(p)), passed in mouse 1 (variant A-9) cells.

The parvoviridae family contains two groups that infect mammalian hosts: (i) defective (helper-dependent) adeno-associated viruses, and (ii) autonomous (helper-independent) parvoviruses. MVM is a member of the latter group. Both groups have been demonstrated to package both plus and minus strands (in separate particles) of the ss-DNA genome, though the minus strand is more typically packaged in the latter group.

The sequence below corresponds to the plus (+) strand, also referred to as the C-strand. The minus (-) strand is also referred to as the V-strand.

The 3' and 5' termini both exhibit the potential for forming stable 'fold-back' hairpins; these sequences appear to play a role in replication [1].

The left and right halves of the genome encode two distinct, but overlapping transcriptional units. The transcripts can be summarized [1] (1 map unit (mu) = 51 bp):

R1 (4.8 kb): 4.5 mu - 46 mu; 46+ mu - 95 mu

R2 (3.3 kb): 4.5 mu - 10.7 mu; 38 mu - 46 mu; 46+ mu - 95 mu

R3 (3.0 kb): 40 mu - 46 mu; 46+ mu - 95 mu

R3 is the major transcript.

There are two major open reading frames, both on the plus (or C) strand. The left side ORF (261-2279) probably encodes a non-capsid protein of 85 kd; the right side ORF probably encodes the viral capsid proteins, VP1 (or A, 83 kd), VP2 (or B, 64 kd), and VP3 (or C, 61 kd). But because of uncertainties about the precise splice points in the transcripts, the exact starts, stops and (possible) intron boundaries are not known.

revision 4804 4870 a-65bp-a in [2]; aa in [1] [2]

revises [1].

[3] sites; splice sites.

FEATURES
source

1. 5149
/organism="Mice minute virus"
/mol_type="genomic DNA"
/strain="MVM(p)"
/db_xref="taxon:10794"
/lab_host="mouse 1 (variant A-9) cell"
114..2279
/gene="NS1"
114..2279
/gene="NS1"
/notes="putative"
/codon_start=1
/product="nonstructural protein"
/protein_id="AAA67108.1"
/db_xref="GI:825481"
/translations="MISGSGSLNQAKRWAPKVKYKOLLKSVYLLFFHSVSRDAQOE
SNOLTWAGNAYSDEVILGATNWLKEKSNQEVFSFVKENVQLNGKDIGHNSYKKELOE
DELKSLQRGNETTWDQSEDMEWETTVDMTKKQVFI PDSLVKKCLFEVLTKNIFPGD
VNVFQHQMGDQCHGVLIIGKQFSDQAGKRWKQVYKOLLKSVYLLFFHSVSRDAQOE
RIKLEADNEQWVLLTYKHKQKDYTKCVLGNVYKVKCHAI CCVLNROGKENTVLFHGA
FLSSDSGKTNFLKEGRLHVSPLYDDMRPETVTTTAQETKGRITKKEVSIK
TTLKELVKRVTSPEWMMQPSYIEMMAQPGGELLKNTLEICTLTARTKTAFLD
ILKAEKSLTNELSPDTRCTIFAFHGNVYKVKCHAI CCVLNROGKENTVLFHGA
STGKSLIAQAQAVNGVCYNANVFPNDCTNKNLIWVEAGNFGQGVNQFKAIC
SGTIRIDQKGGKQIEPTPTVMTNENITVVRIGCEERPEHTQPIRDMNLHLTH
TLFGDFGLVDKNSWPMICAWLVKNGYQSTWASYCAKWKVPDWSNNAEPKPTPTNL
LGSARFPFTPKSPYALTPLASDLELALPEWSTPTNPVAGTAEPTNTGTSAGS
KACQDQLSPTWSEIEEDLRACFGABPLKKDFSEPLNLD"

gene

CDS

NSYKKELODELKSLQRGNETTWDQSEDMEWETTVDMTKKQVFI PDSLVKKCLFEVL
NTKNI PPGDVNMFQWENKQDCHGVLIIGKQFSDQAGKRWKQVYKOLLKSVYLLFFH
CNVQLTPAEIRIKLEADNEQWVLLTYKHKQKDYTKCVLGNVYKVKCHAI CCVLNRO
TSFPRDGGFLSSDSGKTNFLKEGRLHVSPLYDDMRPETVTTTAQETKGRITKKEVSI
QRTKKEVSIKTLNELSPDTRCTIFAFHGNVYKVKCHAI CCVLNROGKENTVLFHGA
ARTKTAFLDILKAEKSLTNELSPDTRCTIFAFHGNVYKVKCHAI CCVLNROGK
NTVLFHGAISGOTIRIDQKGGKQIEPTPTVMTNENITVVRIGCEERPEHTQPIRDM
OVNQFKAICSGTIRIDQKGGKQIEPTPTVMTNENITVVRIGCEERPEHTQPIRDM
RMLNHLTHLTPQDGLVDKNSWPMICAWLVKNGYQSTWASYCAKWKVPDWSNNAE
PKPTPTNLGARSARFPFTPKSPYALTPLASDLELALPEWSTPTNPVAGTAEPTNTG
TNTGTSAGS KACQDQLSPTWSEIEEDLRACFGABPLKKDFSEPLNLD"
2002..2398
/gene="VP"
2002..2280
/gene="VP"
/notes="major transcription start site"
/number=1
2006..2280
/gene="VP"
/notes="minor transcription start site"
/number=1
2009..2280
/gene="VP"
/notes="minor transcription start site"
/number=1
2281..2398
/gene="VP"
/notes="alternative intron"
2281..2376
/gene="VP"
/notes="alternative intron"
join(2286..2316,2399..4557)
/gene="VP1"
join(2286..2316,2399..4557)
/gene="VP1"
/codon_start=1
/protein_id="AAA67111.1"
/db_xref="GI:332294"
/translation="MAPPKAKRKGVPVPGYKYLPGNSLDQGETNPDSAAAKHDE
AYDOYKSKNPVLYESAADOFIDOTKADMGKGVHYFERTKRAFPKLTATSE
GTSVSRAGKTRTPPAYIFINQARAKKLTSSAAQSSQTSMDSGTSQPSDGNVHISAA
RVERAADPGGSGGGGGVGVSTYDNTQHTYRFLGDGWEIITATLRLVHLNMP
KSENYCRIRVHNTDTSYKGNMAKDDAHEQIWPMSLDVACMVAVDSNNILPYTPAA
MSQLNLSLDQEIFNVLTQTEQDLGGQAIKYNNDLTACMVAVDSNNILPYTPAA
NSMETLGFYWPVKATSPYRYFCVDRDLSTYENQEGVHEHNVMTKGMNSQPF
ENTQOITLLRTGDEPATGYTYFDTNSVKLTHTQNTROLGQPLLSTPEADTGLT
TAQSGRHGTTQMGVNVWSEAIRTPAQVGFQCPNDPEASRAGPAAAPKVPADITQGV
DKEANGSVRYSGKHQGENWASHGHPAPERIYTWDETSFGSDTKDGFQISAPLVVPPP
LNGILTNANPIGTKNDIHFSNVFSYGLTAPSHSPSPYPOQIWDKELDLHKKPLRH
ITAPFVCKNNAFGQMLVRLGNPLTDQYDPNGATLSRI VTYGTFFWKGLTWRAKLRAN
TTWNPVYQVSAEDNGNSYMTKWLPTATGNNQSVPLITRTPVARNTY"
2286..2354
/gene="VP"
/notes="ORF1; putative"
/codon_start=1
/product="unknown protein"
/protein_id="AAA67110.1"
/db_xref="GI:825482"
/translation="MAPPKAKRKGKGLRDGLVGY"
2286..2316
/gene="VP"
/number=1
2317..2398
/gene="VP"
/notes="VP intron (alt.)"
2332..2361
/gene="VP"
/notes="ORF3; putative"
/codon_start=1
/product="unknown protein"
/protein_id="AAA67112.1"
/db_xref="GI:825483"
/translation="MVGWNGINV"

mRNA

CDS

gene

exon

exon

exon

exon

intron

intron

gene

CDS

CDS

CDS	2354..2398	Db	861	GTTCCTTTTGGAAACATGATTGCTTACTATTTTAACTAAAGAAAATAAGCACTAGT	920
/gene="VP"					
/note="ORF2; putative"					
/codon_start=1					
/product="unknown protein"					
/protein_id="AAA67113.1"					
/db_xref="GI:825484"					
/translation="MFNYLFYRPEITWF"					
2399..54557					
exon	/gene="VP1"	Db	981	AAAGAGGGCGAGCGCCATCTAGTGAGCAAACTATACCTGATGACATCGGCGCCAGAAACG	10400
/number=2					
2794..4557					
CDS	/gene="VP1"	Db	1041	GTGTGAAAACCAAGTAAACCACTGCGCAGGAAACTAAGCGCGGAGAAATTCAAACTAAAAAA	11000
/note="VP2"					
/codon_start=1					
/protein_id="AAA67114.1"					
/db_xref="GI:332296"					
/translation="MSDCTSQPSGNAVHSAVERAADPGSGGGSGGGGVGVST					
GSYDNQTHYRPLGDGWELTALATRLVHLNPKSENYCRIRVHNTDTSVKGNMAKDD					
AHEQIWTWPSLVANAGVWLQPSDWOYICNTMSQLNLVSLDQEIENFVNLKTVTEQDL					
GGQAIKIYNNDLTACMMVAVDNSNLLPYTPAANSMETLGFYPMKFTIAPYRYFCDV					
RDLSTVYENQEGTVEHNMVGTPEKMSQSFTEIENQOITLLRTGDFEATGTYVDFRNS					
VKLTHWNTNROLGQPELLSTPPEADTDAGTLTAQGRHGTLTGOMGVNHWSEAIRTPA					
QVGFCDPHNDFEARAGPAPKVPADITQGVNDKEANGSVYSYKGKHGHNWASHGPA					
PERYTWDETSFGSGRDTKGFQISAPLWPPPLNGILTNANPIGTAKNDIHFNVFNSY					
GLTAPSHSPSPVYPOGOIWDKELDLHKPLRHITAPFVCKNNAPGQMLVRLGNLTDQ					
Query Match 99.9%; Score 2017.4; DB 14; Length 5149;					
Best Local Similarity 100.0%; Pred. No. 0;					
Matches 2018; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					
Qy	1	ATGCGTGGAAATCGTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA	60		
Db	261	ATGCGTGGAAATCGTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA	320		
Qy	61	AGTAACCGAGAGTGTCTCAATTTGTTTTTAAATGAAATCTTCAACTGAATGGAAA	120		
Db	321	AGTAACCGAGAGTGTCTCAATTTGTTTTTAAATGAAATCTTCAACTGAATGGAAA	380		
Qy	121	GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTACAA	180		
Db	381	GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTACAA	440		
Qy	181	CGAGGACGGAATCTACTTGGGACCAAGCGGAGCATGGAATGGGAAACCAAGTGGAT	240		
Db	441	CGAGGACGGAATCTACTTGGGACCAAGCGGAGCATGGAATGGGAAACCAAGTGGAT	500		
Qy	241	GAAATGACCAAAAAGCAAGTATTCATTTTGGATCTTTGGTTAAAAATGTTTATTGAA	300		
Db	501	GAAATGACCAAAAAGCAAGTATTCATTTTGGATCTTTGGTTAAAAATGTTTATTGAA	560		
Qy	301	GTGCTTTAACCAAGGAATATTTCTCGTGATGTTAATTTGGTTGTGCAACATGAATGG	360		
Db	561	GTGCTTTAACCAAGGAATATTTCTCGTGATGTTAATTTGGTTGTGCAACATGAATGG	620		
Qy	361	GGAAAAAGACCAAGCTGGCACTGCCATGTACTAATTTGGAGGAAGGACTTTAGTCAAGCT	420		
Db	621	GGAAAAAGACCAAGCTGGCACTGCCATGTACTAATTTGGAGGAAGGACTTTAGTCAAGCT	680		
Qy	421	CAAGGGAAATGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCC	480		
Db	681	CAAGGGAAATGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCC	740		
Qy	481	TGTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGACAAT	540		
Db	741	TGTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGACAAT	800		
Qy	541	GAGTGGGTACTCTACTTACTTATAAGCATAAGCAAAACCAAAAAAGACTATACCAAGTGT	600		
Db	801	GAGTGGGTACTCTACTTACTTATAAGCATAAGCAAAACCAAAAAAGACTATACCAAGTGT	860		
Qy	601	GTTCCTTTTGGAAACATGATTGCTTACTATTTTAACTAAAGAAATAAGCACTAGT	660		

```
QY 1741 TCACCAATTCAGCACCGAAAGTACGCCTCTCAGCCAGCAACTATGCACTAACTCCACTT 1800
Db 2001 TCACCAATTCAGCACCGAAAGTACGCCTCTCAGCCAGCAACTATGCACTAACTCCACTT 2060
QY 1801 GCATCGGATCTGAGGACCTGGCTTTAGAGCTTTGAGGACACCAAAATATCTCTGTTGCG 1860
Db 2061 GCATCGGATCTGAGGACCTGGCTTTAGAGCTTTGAGGACACCAAAATATCTCTGTTGCG 2120
QY 1861 GGCACTCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGTTCAA 1920
Db 2121 GGCACTCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGTTCAA 2180
QY 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGATTTGAGAGCGTCTTCGGTGGCGAACC 1980
Db 2181 CTGAGCCCAACTTGGTCAGAGATCGAGGAGATTTGAGAGCGTCTTCGGTGGCGAACC 2240
QY 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACCTGACTAA 2019
Db 2241 TTGAAGAAAGACTTCAGCGAGCGCTGAACCTGACTAA 2279

RESULT 5
AX137743
LOCUS
DEFINITION
Sequence 8 from Patent EPI077260.
ACCESSION
AX137743
VERSION
AX137743.1 GI:14273917
KEYWORDS
Mice minute virus
SOURCE
Mice minute virus
ORGANISM
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
AUTHORS
Nuesch,J. and Rommelaere,J.
TITILE
Parvovirus ns1 variants
JOURNAL
Patent: EP 1077260-A 8 21-FEB-2001;
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts
(DE)

FEATURES
source
Location/Qualifiers
1..2019
/organism="Mice minute virus"
/mol_type="unassigned DNA"
/db_xref="taxon:10794"
/notes="Parvovirus NS1 variant"
1..2019
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC39993.1"
/db_xref="GI:14273918"
/translation="MAGNAYSDEVLGATNWLKEKSNQEVFSFVKFNENVQLNGKDIQW
NSYKELQDELKSLQGAETTWQSEDMEWETTVDEMTKQVFI FDSLKKCLFEVL
NTKNI PPGDVNWFVQHEWGKQDGHCHVLIGKDRSQAGKWRRLNLYVWSRWLVA
CNVQLTPABRIKLREIAEDNEWTLTYKHKQTKDYTKCVLFGNMIAYFLTKKIS
TSPROGGFLSSDSKWNFLKEGGRHLVSKLTDMDRPEVETITVTAQETKRGRI
QTKKEVSIKTKLELHKRVTSPEDDMMOPDSYEMMAQPOGENLLKNTLEICTLTL
ARTKFDLILEKAETSKLTFSPDTRACRIPAFHGMNVKVCNLCVLRNQGGR
NTVLFPGASTGSI IAOIAQAVGVNVCNVAANVPNDCTNKNLIWVEAGNFGQ
OVNQKRAI CSGOTIRIDOKGSGSKQIEPTVIMTNENITVVRIGCEERPEHTQIRD
RLNLIHLTLPGDFGLVDKNPMLICAMLVKNYQSTWASVCAMKWKVPDWSNWE
KVPFTPINLIGSARSFPTPKSTPNTALPTPLASDLSDLALEPWSPTENTPVAGTAE
TONTGAGSKACQDQGLSPTWSEIEBDRACFGEPLKKDFSEPLNLD"

ORIGIN
Query Match 99.8%; Score 2015.8; DB 6; Length 2019;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCTGGAAATGCTTACTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA 60
Db 1 ATGGCTGGAAATGCTTACTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA 60
QY 61 AGTAACGAGGAAGTGTCTCATTTGTTTTAAAAATGAAAAATGTTCAACTGAATGAAAA 120
```

```
Db 61 AGTAACGAGGAAGTGTCTCATTTGTTTTAAAAATGAAAAATGTTCAACTGAATGAAAA 120
QY 121 GATATCGGATGGAAATAGTTTACAAAAAGAGCTGCAGAGGACGAGCTGAAATCTTTTACAA 180
Db 121 GATATCGGATGGAAATAGTTTACAAAAAGAGCTGCAGAGGACGAGCTGAAATCTTTTACAA 180
QY 181 CGAGGAGCGGAAACTACTTTGGGACCAAAGCGAGGACATGGAATGGGAAACCCACAGTGGAT 240
Db 181 CGAGGAGCGGAAACTACTTTGGGACCAAAGCGAGGACATGGAATGGGAAACCCACAGTGGAT 240
QY 241 GAAATGACCAAAAAGCAAGTATTCATTTTGGTCTTTGTTTAAAAAATGTTTATTGAA 300
Db 241 GAAATGACCAAAAAGCAAGTATTCATTTTGGTCTTTGTTTAAAAAATGTTTATTGAA 300
QY 301 GTGCTTTAACACAAAGAAATATATTTCTGCTGTGATGTTAAATTTGTTCTGCAACATGAATGG 360
Db 301 GTGCTTTAACACAAAGAAATATATTTCTGCTGTGATGTTAAATTTGTTCTGCAACATGAATGG 360
QY 361 GGAAGAGACCAAGGCTGGCACTGCCATGTAATAATTGGAGGAAAGGACATTTAGTCAAGCT 420
Db 361 GGAAGAGACCAAGGCTGGCACTGCCATGTAATAATTGGAGGAAAGGACATTTAGTCAAGCT 420
QY 421 CAAAGGAAATGTTGGAGAAAGGCACTAAATGTTTACTGGAGCAGATGGTTGTTAAACAGCC 480
Db 421 CAAAGGAAATGTTGGAGAAAGGCACTAAATGTTTACTGGAGCAGATGGTTGTTAAACAGCC 480
QY 481 TGTAAATGTGCAACTAACACACGCTGAAAGAAATTTAAACTAAGAGAAATAGCAGAAAGCAAT 540
Db 481 TGTAAATGTGCAACTAACACACGCTGAAAGAAATTTAAACTAAGAGAAATAGCAGAAAGCAAT 540
QY 541 GAGTGGTTTACTCTACTTTATAGCATAAGCAAAACCAAAAGACATATACCAAGTGT 600
Db 541 GAGTGGTTTACTCTACTTTATAGCATAAGCAAAACCAAAAGACATATACCAAGTGT 600
QY 601 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAATAAGCAGTGT 660
Db 601 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAATAAGCAGTGT 660
QY 661 CCACCAAGAGACGAGGCTATTTCTTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA 720
Db 661 CCACCAAGAGACGAGGCTATTTCTTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA 720
QY 721 AAAGAGGCGAGCGGCATCTAGTGAGCAAACTATACACTGATGACATCGCGGCAGAAACG 780
Db 721 AAAGAGGCGAGCGGCATCTAGTGAGCAAACTATACACTGATGACATCGCGGCAGAAACG 780
QY 781 GTTGAACCAACAGTAAACCACTCGCGAGGAACTTAAGCGCGGAGAACTCAAACTAAAAA 840
Db 781 GTTGAACCAACAGTAAACCACTCGCGAGGAACTTAAGCGCGGAGAACTCAAACTAAAAA 840
QY 841 GAAAGTTGCTATTAAAACTACACTTAAAGAGCTGGTGCATAAAAGAGTAAACCTCACCAGAG 900
Db 841 GAAAGTTGCTATTAAAACTACACTTAAAGAGCTGGTGCATAAAAGAGTAAACCTCACCAGAG 900
QY 901 GACTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Db 901 GACTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 961 AACCTGCTGAAAAATACCGTAGAGATTTGTACACTCACTTAGCCAGACCAAAACAGCA 1020
Db 961 AACCTGCTGAAAAATACCGTAGAGATTTGTACACTCACTTAGCCAGACCAAAACAGCA 1020
QY 1021 TTTGACTTAAATTTAGAAAAAGCTGAAACCAAGCAAACTAAACCACTTTTCACTGCTGAC 1080
Db 1021 TTTGACTTAAATTTAGAAAAAGCTGAAACCAAGCAAACTAAACCACTTTTCACTGCTGAC 1080
QY 1081 ACAAGAACCTGAGAAATTTTCTGCTTTTATGCTGGAATATGTTAAAGTTTGCATGCT 1140
Db 1081 ACAAGAACCTGAGAAATTTTCTGCTTTTATGCTGGAATATGTTAAAGTTTGCATGCT 1140
QY 1141 ATTTGCTGTGTTTTTAAACAGACAAAGGCGCAAAAGAAATACCTGTTTATTTTCAATGACCA 1200
Db 1141 ATTTGCTGTGTTTTTAAACAGACAAAGGCGCAAAAGAAATACCTGTTTATTTTCAATGACCA 1200
```


QY	1201	GCAGCAGCAGCAATCTATTATTCACAGCCATAGCACAGCAGTTGGCAATGTTGGT	1260
DB	1201	GCAGCAGCAGCAGCAATCTATTATTCACAGCCATAGCACAGCAGTTGGCAATGTTGGT	1260
QY	1261	TGCTATAATGACGCAATGTAATCTTCCATTTAATGACTGTACCAACAGAACTTGATT	1320
DB	1261	TGCTATAATGACGCAATGTAATCTTCCATTTAATGACTGTACCAACAGAACTTGATT	1320
QY	1321	TGGGTAGAAGAGCTGGTAACTTTGGACAGCAAGTAACAGATTTAAAGCCATTGCTCT	1380
DB	1321	TGGGTAGAAGAGCTGGTAACTTTGGACAGCAAGTAACAGATTTAAAGCCATTGCTCT	1380
QY	1381	GGTCAAACTATTGCGATTGATCAAAAGGAAAGGAGCAAGCAAGATTTGAACCAACCA	1440
DB	1381	GGTCAAACTATTGCGATTGATCAAAAGGAAAGGAGCAAGCAAGATTTGAACCAACCA	1440
QY	1441	GTCTATCATGACCAATGAGACATTTACAGTGTGTGAGATAGCTGCGAAGAAAGACCA	1500
DB	1441	GTCTATCATGACCAATGAGACATTTACAGTGTGTGAGATAGCTGCGAAGAAAGACCA	1500
QY	1501	GAACACACTCAACCAATTCAGAGACAGAAATGCTTTAAACATTTCACTAACACATCTTGCCT	1560
DB	1501	GAACACACTCAACCAATTCAGAGACAGAAATGCTTTAAACATTTCACTAACACATCTTGCCT	1560
QY	1561	GGTGACTTTGGTTGGTTGACAAAATGAATGGCCCAATGATTTGCTTGGTTGGTAAAG	1620
DB	1561	GGTGACTTTGGTTGGTTGACAAAATGAATGGCCCAATGATTTGCTTGGTTGGTAAAG	1620
QY	1621	AATGGTTTACCAATCTACATGCGAAGCTACTGTCTAAATGGGCAAGTTCTCTGATTGG	1680
DB	1621	AATGGTTTACCAATCTACATGCGAAGCTACTGTCTAAATGGGCAAGTTCTCTGATTGG	1680
QY	1681	TCAGAAAACCTGGGCGGAGCCAAAGGTGCGCAACTCTATAAATTTACTAGTTGCGCACGC	1740
DB	1681	TCAGAAAACCTGGGCGGAGCCAAAGGTGCGCAACTCTATAAATTTACTAGTTGCGCACGC	1740
QY	1741	TCACCAATTCAGACACCGAAAGATGACGCTCTCAGCGAGCAATATGCACTAACCTCCACTT	1800
DB	1741	TCACCAATTCAGACACCGAAAGATGACGCTCTCAGCGAGCAATATGCACTAACCTCCACTT	1800
QY	1801	GCACTCGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACCAATATCTCTCTGTTGG	1860
DB	1801	GCACTCGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACCAATATCTCTCTGTTGG	1860
QY	1861	GGCACTGCGAGAAACCCAGAACACTTGGGAAAGCTGGTTCCAAAGCTGCGCAAGATGGTCAA	1920
DB	1861	GGCACTGCGAGAAACCCAGAACACTTGGGAAAGCTGGTTCCAAAGCTGCGCAAGATGGTCAA	1920
QY	1921	CTGAGCCCAACTTGGTTCAGAGATCGAGGAGATTTGAGAGCGTGTCTCGGTGCGGAACCG	1980
DB	1921	CTGAGCCCAACTTGGTTCAGAGATCGAGGAGATTTGAGAGCGTGTCTCGGTGCGGAACCG	1980
QY	1981	TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA	2019
DB	1981	TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA	2019

RESULT	6
AX137747	
LOCUS	AX137747 2019 bp DNA linear PAT 30-MAY-2001
DEFINITION	Sequence 12 from Patent EP1077260.
ACCESSION	AX137747
VERSION	AX137747.1 GI:14273921
KEYWORDS	Mice minute virus
SOURCE	Mice minute virus
ORGANISM	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE	1
AUTHORS	Nuesch, J. and Rommelaere, J.
TITLE	Parvovirus ns1 variants
JOURNAL	Patent: EP 1077260-A 12 21-FEB-2001;
	Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts

(DE)

FEATURES	Location/Qualifiers
source	1. .2019
	/organism="Mice minute virus"
	/mol_type="unassigned DNA"
	/db_xref="taxon:10794"
	/note="Parvovirus NS1 variant"
CDS	1. .2019
	/note="unnamed protein product"
	/codon_start=1
	/protein_id="CAC39995.1"
	/db_xref="GI:14273922"
	/translation="MAGNAYSDVGLGATNWLKEKSNQEVFSFVKENVOLNGKDIGM
	NSYKELQDELKSLQRGAEETWDQSDMEWETTVDEMFKQVFIEDSLVKKCLFEVL
	NTKIPPGDVNMFQHEWKGDKGHWLIGGKDFSOAQKQWRRQLNVYFWSRWLVTA
	CNVQLTPAERIKLRETAEDNWTLLTYKHQTKDYTKCVLFENMIAYVFLPKKIS
	TSPRGGYVFLSSDSQKTNFLKGERHLVSKLYTDDMRPETVETVTTTAQETKGR
	OTKQVSIYKTLTKELVHKVTPEDMMWQPDYSIEMMAQPGGKLNKTLKLEICTLTL
	ARTKTAFDLLEKAETSKLTNPSLPTDTRCIRFAFGHMYVVKYCHALCCVLNRQGGK
	NAVLFPGASTGKSIQAALAAQAGNVGCTNNAANVFPNDCTNKLILWVEAGNGGQ
	QVNFKAICSGQIRIDQKQSGKQIEPTFVIMTNNITVVRIGCEERPEHTOPIRD
	RMLNHLTLPFGDLGVDKNWPMICAWLVKNGYQSTWASVCAKWKVDPDWSNNAE
	PKVPTPINLLSGARSPTPKSTPLSQNTALPLASDLLEDALEPMTNPVAGTAE
	TQNTGEAGSKACQDGLSPTWSEIEEDLRACFGAEPLKKDFSEPLND"

Query Match 99.8%; Score 2015.8; DB 6; Length 2019;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	ATGGCTGGAAATGCTTACTCTGATGAAGTTTTCGGAGCAACCACTGGTTAAAGGAAAA	60
DB	1	ATGGCTGGAAATGCTTACTCTGATGAAGTTTTCGGAGCAACCACTGGTTAAAGGAAAA	60
QY	61	AGTAACCAAGCAAGTGTCTCAATTTGTTTAAATAATGAAATGTTCACCTGAATGGA	120
DB	61	AGTAACCAAGCAAGTGTCTCAATTTGTTTAAATAATGAAATGTTCACCTGAATGGA	120
QY	121	GATATCGGATGGAATAGTTTACAAAAAGAGCTCGAGAGGACGAGCTGAAATCTTTA	180
DB	121	GATATCGGATGGAATAGTTTACAAAAAGAGCTCGAGAGGACGAGCTGAAATCTTTA	180
QY	181	CGAGGAGCGGAACTACTTGGGACCAAGCGAGGACATGGAATGGGAAACCACTGGAT	240
DB	181	CGAGGAGCGGAACTACTTGGGACCAAGCGAGGACATGGAATGGGAAACCACTGGAT	240
QY	241	GAAATGACCAAAAAAGCAAGTATTCTTTTGGTTTAAATAATGTTTATTTGAA	300
DB	241	GAAATGACCAAAAAAGCAAGTATTCTTTTGGTTTAAATAATGTTTATTTGAA	300
QY	301	GTGCTTTAACAACAAAGAAATATATTCCTGCTGATGTTAAATTTGTTGCAACATGA	360
DB	301	GTGCTTTAACAACAAAGAAATATATTCCTGCTGATGTTAAATTTGTTGCAACATGA	360
QY	361	GGAAAAAGCAAGCTGCGCACTGCCATGTAATTTGGAGGAAAGGACTTTAGTCAAG	420
DB	361	GGAAAAAGCAAGCTGCGCACTGCCATGTAATTTGGAGGAAAGGACTTTAGTCAAG	420
QY	421	CAAGGGAATGTTGGAGAGAGGCAACTAAATTTTACTTGGAGCAGATGCTTGGTAA	480
DB	421	CAAGGGAATGTTGGAGAGAGGCAACTAAATTTTACTTGGAGCAGATGCTTGGTAA	480
QY	481	TGTAATGTGCAACTAAACACAGCTGAAAGAAATTTAAACTAAGAGAAATAGCAGA	540
DB	481	TGTAATGTGCAACTAAACACAGCTGAAAGAAATTTAAACTAAGAGAAATAGCAGA	540
QY	541	GAGTGGGTTACTCTTACTTATAGCATAACCAACCAAAAAAGACTTATACCAAGT	600
DB	541	GAGTGGGTTACTCTTACTTATAGCATAACCAACCAAAAAAGACTTATACCAAGT	600
QY	601	GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAAAAAGCACTAG	660
DB	601	GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAAAAAGCACTAG	660

661 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA 720
Db
661 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA 720
QY 721 AAGAAGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCGAGAAACG 780
Db 721 AAGAAGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCGAGAAACG 780
QY 781 GTTGAACACACAGTAACCTGCGCAGGAACTTAAGCGCGCAGAACTTCAAACTAAAAA 840
Db 781 GTTGAACACACAGTAACCTGCGCAGGAACTTAAGCGCGCAGAACTTCAAACTAAAAA 840
QY 841 GAAGTTGCTATTAAAACTACATTTAAAGAGCTGGTGATATAAAGAGTAACCTCACCAGAG 900
Db 841 GAAGTTTCTATTAAAACTACATTTAAAGAGCTGGTGATATAAAGAGTAACCTCACCAGAG 900
QY 901 GACTGATGATGACGACGACAGTACATTTGAATGATGCTCAACAGGTGGAGAA 960
Db 901 GACTGATGATGACGACGACAGTACATTTGAATGATGCTCAACAGGTGGAGAA 960
QY 961 AACTGCTGAAAAATACGCTAGAGATTGTACACTAACTAGCCAGAACCAAAACAGCA 1020
Db 961 AACTGCTGAAAAATACGCTAGAGATTGTACACTAACTAGCCAGAACCAAAACAGCA 1020
QY 1021 TTTGACTTAATTTTAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCTGAC 1080
Db 1021 TTTGACTTAATTTTAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCTGAC 1080
QY 1081 ACAAGAACCTGCAAGATTTTGGCTTTTCATGGCTGGAACATATGTTAAAGTTTGCATGCT 1140
Db 1081 ACAAGAACCTGCAAGATTTTGGCTTTTCATGGCTGGAACATATGTTAAAGTTTGCATGCT 1140
QY 1141 ATTGCTGCTTTTAAACACAGAGGACCAAGAAATCTGTTTATTTTCAATGACCA 1200
Db 1141 ATTGCTGCTTTTAAACACAGAGGACCAAGAAATCTGTTTATTTTCAATGACCA 1200
QY 1201 GCCAGCACAGGCAATCTATTATTGACCAAGCCATAGCAACAGTGGCAATGTTGGT 1260
Db 1201 GCCAGCACAGGCAATCTATTATTGACCAAGCCATAGCAACAGTGGCAATGTTGGT 1260
QY 1261 TGCTATAATGACGCAATGAACTTCCATTTAATGACTGTACCAACAGAACTTGAT 1320
Db 1261 TGCTATAATGACGCAATGAACTTCCATTTAATGACTGTACCAACAGAACTTGAT 1320
QY 1321 TGGGTAGAAGAGCTGTAATTTTGACAGCAAGTAAACAGTTTAAAGCAATTTGCTCT 1380
Db 1321 TGGGTAGAAGAGCTGTAATTTTGACAGCAAGTAAACAGTTTAAAGCAATTTGCTCT 1380
QY 1381 GGTCAAACTATTGCAATGATCAAAAGGAAAGGAGCAAAACAGATTGAACCAACCA 1440
Db 1381 GGTCAAACTATTGCAATGATCAAAAGGAAAGGAGCAAAACAGATTGAACCAACCA 1440
QY 1441 GTCATATGACCAAAATGAGAACATACAGTGGTCAAGATAGGCTGCGAAGAAAGCA 1500
Db 1441 GTCATATGACCAAAATGAGAACATACAGTGGTCAAGATAGGCTGCGAAGAAAGCA 1500
QY 1501 GAACACACTCAACCAATCAGAGACAGATGCTTAAACATTCATCTAACACATCTTGCT 1560
Db 1501 GAACACACTCAACCAATCAGAGACAGATGCTTAAACATTCATCTAACACATCTTGCT 1560
QY 1561 GGTGACTTTGGTTGGTTGA CAAAAATGAATGGCCCATGATTGCTGTTGGTTGTAAG 1620
Db 1561 GGTGACTTTGGTTGGTTGA CAAAAATGAATGGCCCATGATTGCTGTTGGTTGTAAG 1620
QY 1621 AATGGTTACCAATCTACCATGGCAAGCTACTGCTAAATGGGGCAAAAGTTCTGATTGG 1680
Db 1621 AATGGTTACCAATCTACCATGGCAAGCTACTGCTAAATGGGGCAAAAGTTCTGATTGG 1680
QY 1681 TCAGAAAACTGGGCGGAGCCAAAGGTGCCAATCTCTATAAATTTACTAGTTTCGGCAGC 1740
Db 1681 TCAGAAAACTGGGCGGAGCCAAAGGTGCCAATCTCTATAAATTTACTAGTTTCGGCAGC 1740

1741 TCACCAATTCAGACACCGAAAAAGTAGCGCTCTCAGCCAGCAACTATGCACTAACTCCACTT 1800
Db 1741 TCACCAATTCAGACACCGAAAAAGTAGCGCTCTCAGCCAGCAACTATGCACTAACTCCACTT 1800
QY 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGAGCACACCAAAATACTCTGTTGG 1860
Db 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGAGCACACCAAAATACTCTGTTGG 1860
QY 1861 GGCACCTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCCAAAGATGTCAC 1920
Db 1861 GGCACCTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCCAAAGATGTCAC 1920
QY 1921 CTGAGCCCCAACTTGGTTCAGAGATCGAGAGGATTTGAGAGCGTGTCTTCGGTGGGAACCG 1980
Db 1921 CTGAGCCCCAACTTGGTTCAGAGATCGAGAGGATTTGAGAGCGTGTCTTCGGTGGGAACCG 1980
QY 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2019
Db 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2019

RESULT 7
AX137751
LOCUS AX137751 2019 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 16 from Patent EP1077260.
ACCESSION AX137751
VERSION AX137751.1 GI:14273925
KEYWORDS
SOURCE Mice minute virus
ORGANISM Mice minute virus
VIRUSES; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE 1
AUTHORS Nueesch, J. and Rommelaere, J.
TITLE Parvovirus ns1 variants
JOURNAL Patent: EP 1077260-A 16 21-FEB-2001;
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts (DE)

FEATURES
source
1. .2019
/organism="Mice minute virus"
/mol_type="unassigned DNA"
/db_xref="taxon:10794"
/note="Parvovirus NS1 variant"
1. .2019
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC39997.1"
/translation="MAGNAYSDEVILGATNWLKESNQEVFVFKNENVOLNGKDIGW
NSYKLEQDELKSLQGAETTWDSQSEDMETTTVDMTKKQVFI FDSLVRKCLFEVL
NTKNI FPGDVNMFQHEWKGDKQCHVLIGGKDFSOAQKWRQLNLYWRSWLVT
CNVOLTPAERI KLREIAEDNEWVLLTVKHOTKKDYTKCVLFGNMIA YVFLTKKIS
TSPPDGGYFLSSDSGKNTNFKEGEHLVSKLYTDDMRPETVETVTTAOSTKGR
QTKSVSINTLKELVHRKVTSPEDMMQPSYIEMMQPGENLLKNTLSICTLTL
ARTKAFDLILEKETSALTNFSLPTRTCRIFAFHGMNVKVCCHAICCLVLRQGGK
NTVLFHGPASTGKSI IAAIAQAVNGVCTNAANVPNDCTNKNLIWVBAGNFGQ
QMNOKAICSGOARIDQKGSKOIETPTVIMTNENITVVRIGCEERPEHTQPTRD
RMLNTHLTHLPGDFGLVDKNEWPMI CAWLKNGVQSTWASCAWKGVKPDNSENWAE
PKVPTPINLIGASRSPFTPKSTPLSONVALTPLASDLLEDALEPWSTPNTPVAGTAE
TQNTGEAGSKACQDQGLSPTWSEIBEDLRCFCAEPLKKDFSEPLND"

CDS
Query Match 99.8%; Score 2015.8; DB 6; Length 2019;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

ORIGIN
QY 1 ATGGCTGGAATGCTTACTCTGATGAACTTTGGGAGCAACCAACTGGTTAAGGAAAA 60
Db 1 ATGGCTGGAATGCTTACTCTGATGAACTTTGGGAGCAACCAACTGGTTAAGGAAAA 60
QY 61 AGTAACCCAGGAGTCTTCTCAATTTTAAAAATGAAAAATGTTCACTGAATGAAAA 120
Db 61 AGTAACCCAGGAGTCTTCTCAATTTTAAAAATGAAAAATGTTCACTGAATGAAAA 120

121 GATATCGGATGGAATAGTTACAAAAAGAGCTGCGAGAGAGCGAGCTGAAATCTTTACAA 180
121 GATATCGGATGGAATAGTTACAAAAAGAGCTGCGAGAGAGCGAGCTGAAATCTTTACAA 180
181 CGAGGAGCGGAACTACTCTGGGACCAAGCGAGGACATGGAATGGGAAACCCACAGTGGAT 240
181 CGAGGAGCGGAACTACTCTGGGACCAAGCGAGGACATGGAATGGGAAACCCACAGTGGAT 240
241 GAAATGACAAAAAGCAAGATTAATTTTGAATCTTTTGGTAAAGAAATGTTTATTGAA 300
241 GAAATGACAAAAAGCAAGATTAATTTTGAATCTTTTGGTAAAGAAATGTTTATTGAA 300
301 GTGCTTTAACCAAGAATATATTTCTGGTGATGTTAAATTTGGTTGTGCAACATGAATGG 360
301 GTGCTTTAACCAAGAATATATTTCTGGTGATGTTAAATTTGGTTGTGCAACATGAATGG 360
361 GGAAGAGACCAAGCTGGCACTGCATGTACTTAATTTGGAGAAAGGACTTTTAGTCAAGCT 420
361 GGAAGAGACCAAGCTGGCACTGCATGTACTTAATTTGGAGAAAGGACTTTTAGTCAAGCT 420
421 CAAGGGAAATGTTGGAGAAAGCACTAAATGTTTACTGGAGCAGATGTTGGTAACAGCC 480
421 CAAGGGAAATGTTGGAGAAAGCACTAAATGTTTACTGGAGCAGATGTTGGTAACAGCC 480
481 TGTAAATGTGCAACTAACACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT 540
481 TGTAAATGTGCAACTAACACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT 540
541 GAGTGGTGTACTTACTTACTTACTTAAAGCAATAGCAAAACCAAAAGAAAGTATACCAAGTGT 600
541 GAGTGGTGTACTTACTTACTTACTTAAAGCAATAGCAAAACCAAAAGAAAGTATACCAAGTGT 600
601 GTTCTTTTGGAAACATGATGTTGCTTACTATTTTAACTAAAGAAAGTAAAGCACTAGT 660
601 GTTCTTTTGGAAACATGATGTTGCTTACTATTTTAACTAAAGAAAGTAAAGCACTAGT 660
661 CCAACCAAGAGACGAGGCTATTTTCTTAGCAGTGAATCTGCTGGGAAAGTAACTAATCTTTTA 720
661 CCAACCAAGAGACGAGGCTATTTTCTTAGCAGTGAATCTGCTGGGAAAGTAACTAATCTTTTA 720
721 AAGAGAGGCGAGCGCACTAGTGAAGCAATACACTGATGACATGCGGCGCAGAAACG 780
721 AAGAGAGGCGAGCGCACTAGTGAAGCAATACACTGATGACATGCGGCGCAGAAACG 780
781 GTTGAACCAACAGTAACTGCGCAGGAACTAAGCGCGGAGAAATTTCAAACTAAAGAA 840
781 GTTGAACCAACAGTAACTGCGCAGGAACTAAGCGCGGAGAAATTTCAAACTAAAGAA 840
841 GAAATGCTATTAAACTACACTTAAAGAGCTGGTGCAATAAAGAGTAACTCAACAGAG 900
841 GAAATGCTATTAAACTACACTTAAAGAGCTGGTGCAATAAAGAGTAACTCAACAGAG 900
901 GACTGGATGATGACGAGCAGACAGTTTACATTTGAATGATGCTCAACAGCTGGAGAA 960
901 GACTGGATGATGACGAGCAGACAGTTTACATTTGAATGATGCTCAACAGCTGGAGAA 960
961 AACTGCTGAAAAATACGCTAGAGATTTGTACACTACTCTAGCCAGAAACCAAAACAGCA 1020
961 AACTGCTGAAAAATACGCTAGAGATTTGTACACTACTCTAGCCAGAAACCAAAACAGCA 1020
1021 TTTGACTTAATTTTAAAAAGCTGAAACCCAGCAAACTAACCACTTTTCACTGCCTGAC 1080
1021 TTTGACTTAATTTTAAAAAGCTGAAACCCAGCAAACTAACCACTTTTCACTGCCTGAC 1080
1081 ACAAGAACCTGCAGAAATTTTCTTTTATGCTGGTGAACATGTTAAAGTTTGCATGCT 1140
1081 ACAAGAACCTGCAGAAATTTTCTTTTATGCTGGTGAACATGTTAAAGTTTGCATGCT 1140
1141 ATTTGCTGTGTTTAAACAGACAGGAGGCAAGAAATACTGTTTATTTTATTCATGACCA 1200
1141 ATTTGCTGTGTTTAAACAGACAGGAGGCAAGAAATACTGTTTATTTTATTCATGACCA 1200

1201 GCAGACAGGCAAAATCTATTATTCAGCAAGCCATAGCACAAGCAGTTGGCAATGTTGGT 1260
1201 GCAGACAGGCAAAATCTATTATTCAGCAAGCCATAGCACAAGCAGTTGGCAATGTTGGT 1260
1261 TGCTATAATGCGCCCAATGTAAACTTTTCCATTTAAATGACTGTACCAACCAAGAACTTGAT 1320
1261 TGCTATAATGCGCCCAATGTAAACTTTTCCATTTAAATGACTGTACCAACCAAGAACTTGAT 1320
1321 TGGGTAGAGAAAGCTGGTAACTTTTGGACAGCAAGTAAACCAAGTTTAAAGCAATTTGCTCT 1380
1321 TGGGTAGAGAAAGCTGGTAACTTTTGGACAGCAAGTAAACCAAGTTTAAAGCAATTTGCTCT 1380
1381 GGTCAAATCTATTCGCAATGATCAAAAGGAAAGGAGGAGCAAGATTTGAACCAACACCA 1440
1381 GGTCAAAGCTATTCGCAATGATCAAAAGGAAAGGAGGAGCAAGATTTGAACCAACACCA 1440
1441 GTCATCATGACCAAAATGAGAACTTACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA 1500
1441 GTCATCATGACCAAAATGAGAACTTACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA 1500
1501 GAACACACTCAACCAATCGAGACAGAAATGCTTAAACATTCATCTAAACATACCTTGCCT 1560
1501 GAACACACTCAACCAATCGAGACAGAAATGCTTAAACATTCATCTAAACATACCTTGCCT 1560
1561 GGTGACTTTGGTTGGTTGACAAAATGATGGCCCATGATTTGTGCTTGGTGGTAAAG 1620
1561 GGTGACTTTGGTTGGTTGACAAAATGATGGCCCATGATTTGTGCTTGGTGGTAAAG 1620
1621 AATGGTTTACCATCTACCATGGCAAGCTACTGTCTTAAATGGGGCAAGTTTCTGTATGG 1680
1621 AATGGTTTACCATCTACCATGGCAAGCTACTGTCTTAAATGGGGCAAGTTTCTGTATGG 1680
1681 TCAGAAAACTGGCGGAGCCAAAGTGGCCAACTCTTATAAATTTACTAGTTCGGCAGC 1740
1681 TCAGAAAACTGGCGGAGCCAAAGTGGCCAACTCTTATAAATTTACTAGTTCGGCAGC 1740
1741 TCACCAATTCACGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTTAACCTCACTT 1800
1741 TCACCAATTCACGACACCGAAAGTACGCTCTCAGCCGAACTATGCACTTAACCTCACTT 1800
1801 GCATCGATCTCGAGGACCTGGCTTTTAGAGCTTTGGAGCACACCAATACTCTCTGTTGCG 1860
1801 GCATCGATCTCGAGGACCTGGCTTTTAGAGCTTTGGAGCACACCAATACTCTCTGTTGCG 1860
1861 GGCACTGCAGAAACCCAGAACTGCGGGAAGCTGGTTCCAAAGCTGCCAAGATGGTCAA 1920
1861 GGCACTGCAGAAACCCAGAACTGCGGGAAGCTGGTTCCAAAGCTGCCAAGATGGTCAA 1920
1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGGTGGGAAACCG 1980
1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGGTGGGAAACCG 1980
1981 TTCAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019
1981 TTCAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019

RESULT 8
MWICG 5085 bp ss-DNA linear VRL 05-JUL-1995
LOCUS Minute virus of mice (MVM(i)), a lymphotropic variant of MVM,
DEFINITION complete genome.
ACCESSION M12032
VERSION M12032.1 GI:332289
KEYWORDS alternative splicing; capsid protein; complete genome;
nonstructural protein.
SOURCE Mice minute virus
ORGANISM Mice minute virus
REFERENCE 1 (bases 1 to 5085)
AUTHORS Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
Tatell,C.R., Gardiner,E.M. and Tattersall,P.
TITLE DNA sequence of the lymphotropic variant of minute virus of mice,
MVM(i), and comparison with the DNA sequence of the fibrotropic

JOURNAL MEDLINE COMMENT	prototype strain J. Virol. 570, 656-669 (1986) 86115415 Original MVM) DNA, clone pEG222.	source text: Minute virus of mice (lymphotropic variant of MVM) DNA, clone pEG222.	Location/Qualifiers	gene	CDS	/gene="vp" /note="alternative intron" join(2287..2317,2400..4558) /gene="vp1" join(2287..2317,2400..4558) /gene="vp1" /codon_start=1 /protein_id="AAA69569.1" /db_xref="GI:332290" /translation="MAPPAAKRAKRWVPGYKYLPGNSLDQGEPTNPSPDAAAKHEDE AYDOYIKSGKNPLYFSAADQRIQDTKADKMGKGVHYFRTKFAFAPKLTADSE GTSVSRAGKRTRPPAYIFINQARAKKLTSSAAQSSQTSMDGTSPQDGGNVAHSA RVERAADPGSGGGGGGSGVSTGYNDQTHYRFLGDMWEITLATRLVHLNMP KSENYCRIRVHNTDTSVKGNAKDDAHEQIWTWPSLVANAGVWLQPSDQYICNT MSQLNLVSDQEIFNVVLKTVTEQDSGGQAIKIYNNDLTACMVAVDNNILPYTAA NSMETLGFYPWKPTIASPVRYVFCVDRDLSVTYNOEGTIEHNVMTGPKGMSORFTI ENTQIITLLRTGDEPATGYTFDTPNPVKLTHTWQNRQLGQPLLTSTFEADTADGTL TAQGRHGAQTMEVNVVSEAIRTRPAQVGCQPHNDFEASRAGFAAPKVPADVTQGV DREANGSVRYSGKHGHNAAHGPAPERIYTDETNFGSGRDRDGFIOASPLVPPP LNGIILTANPIGTKNDIHFSNVFNSYGPLTAFSHSPVYPOGQIWDKELDLHKPLRH ITAPFVCKNNAPOQLVRLGPLNLTDOYDPNGATLSRIYTYGTFFWKGLTMRKLRAN TTWNPVYQVSVEDNGNSYMSVTKWLPTATGNNQSVPLITRPVARNITY" 2287..2355 /gene="vp1" /note="ORF1; putative" /codon_start=1 /product="unknown protein" /protein_id="AAA69568.1" /db_xref="GI:825478" /translation="MAPPAAKRAKRGKGLRDGLVGY" <2287..2317 /gene="vp1" /number=2 2318..2399 /gene="vp1" /note="alternative intron" 2333..2362 /gene="vp1" /note="ORF3; putative" /codon_start=1 /product="unknown protein" /protein_id="AAA69570.1" /db_xref="GI:825479" /translation="MVGWGWGINV" 2355..2399 /gene="vp1" /note="ORF2; putative" /codon_start=1 /product="unknown protein" /protein_id="AAA69571.1" /db_xref="GI:825480" /translation="MFNYLFYRPEITWF" 2400..44558 /gene="vp1" /number=2 2795..4558 /gene="vp2" 2795..4558 /gene="vp2" /codon_start=1 /protein_id="AAA69572.1" /db_xref="GI:332292" /translation="MSDGTSPQDGGNVAHSAARVERAADPGSGGGGGGGVGYST GSYDQTHYRFLGDMWEITLATRLVHLNMPKSENYCRIRVHNTDTSVKGNAKDD AHEQIWTWPSLVANAGVWLQPSDQYICNTMSQLNLVSDQEIFNVVLKTVTEQDS GGQAIKITWPSLVANAGVWLQPSDQYICNTMSQLNLVSDQEIFNVVL
-------------------------------	--	--	---------------------	------	-----	--

ORIGIN	1085 bp upstream of EcoRI site.	
Query Match	94.4%; Score 1905.4; DB 14; Length 5085;	
Best Local Similarity	96.5%; Pred. No. 0;	
Matches 1948; Conservative	0; Mismatches 71; Indels 0; Gaps 0;	
QY	1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGACCAACCACTGGTTAAAGGAAAAA 60	1021 TTTGACTTAATTTTAGAAAAGCTGAAACCCAGCAAACTAACCACTTTTCTACTGCTGAC 1080
DB	262 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAAACCAACCACTGGTTAAAGGAAAAA 321	1282 TTTGACTTGAATTTTAGAAAAGCTGAAACCCAGCAAACTAACCACTTTTCTACTGCTGAC 1341
QY	61 AGTAACCCAGGAAGTGTCTCAATTTGTTTAAAAATGAAATGTTCAACTGAAATGGAATA 120	1081 ACAAGAACCTGCAAGATTTTGTCTTTTATGCGGTGGAACCTATGTAAAGTTTGGCATGCT 1140
DB	322 AGTAACCCAGGAAGTGTCTCAATTTGTTTAAAAATGAGATGTTCAACTTAATGGAATA 381	1342 ACAAGAACCTGCAAGATTTTGTCTTTTATGCGGTGGAACCTATGTAAAGTTTGGCATGCT 1401
QY	121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGAGAGGACGAGCTGAAATCTTTACAA 180	1141 ATTTGCTGTGTTTAAACAGACAGAGGCAAAAGAAATACTGTTTATTTATTCATGGAACA 1200
DB	382 GATATCGGATGGAATAGTTTACAAAAAGAGCTGAGAGGACGAGCTGAAATCTTTACAA 441	1402 ATTTGCTGTGTTCTAAACAGACAGAGGCAAAAGAAATACTGTTTATTTATTCACGGAACA 1461
QY	181 CGAGGAGCGGAACTACTTGGGACCAAGCGAGGACATGGAATGGGAAACCAACAGTGGAT 240	1201 GCCAGACAGGCAAAATCTATTTTGCACAGCCATAGCAACAGCAGTTGCAATGTTGGT 1260
DB	442 CGAGGAGCGGAACTACTTGGGACCAAGCGAGGACATGGAATGGGAAATCTACAGTGGAT 501	1462 GCCAGTACAGGCAAAATCTATTTTGCACAGCCATAGCAACAGCAGTTGTTAATGTTGGT 1521
QY	241 GAAATGACCAAAAAAGCAAGTATTCAATTTTGAATCTTTGTTTAAAAAATGTTTATTGAA 300	1261 TGTCTAATATGCAAGCAATGTAAACTTTTCAATTTTAAATGACTGTACCAACAGAACTTGATT 1320
DB	502 GAAATGACCAAAAAAGCAAGTATTCAATTTTGAATCTTTTGAATGTTTATTGAA 561	1522 TGTCTAATATGCAAGCAATGTAAACTTTTCAATTTTAAATGACTGTACCAACAGAACTTGATT 1581
QY	301 GTCTTTAAACAAAGAAATATATTTCTGTTGATGTTAATTTGTTTGTGCAACATGAATGG 360	1321 TGGGTAGAAAGAGCTGTTAACTTTTGGACAGCAAGTAAACCAAGTTTAAAGCCATTTGCTCT 1380
DB	562 GTCTTTAGCAAAAAATATAGCTCTGCTGATGTTACTTTGTTTGTGCAAGCATGAATGG 621	1582 TGGGTAGAAAGAGCTGTTAACTTTTGGACAGCAAGTAAACCAAGTTTAAAGCCATTTGCTCT 1641
QY	361 GGAAGACCAAGCTGGCACTGCCATGTAATTAATTTGGAGGAAGGACTTTTACTCAAGCT 420	1381 GGTCAAACTATTTCGCATTTGATCAAAAAGGAAAGGAGCAAGCAAGATTTGAACCAACCA 1440
DB	622 GGAAGACCAAGCTGGCACTGCCATGTAATTAATTTGGAGGAAGGACTTTTACTCAAGCT 681	1642 GGTCAAACTATTTCGCATTTGATCAAAAAGGAAAGGAGCAAGCAAGATTTGAACCAACCA 1701
QY	421 CAAGGCAAAATGTTGGAGAGCAAGCAATTAATTTTACTTGGAGCAGATGTTTGGTAAACAGCC 480	1441 GTCATCATGACCAAAATGAGAACATTTACAGTGGTCAAGATAGGCTGCGAAGAGAGACCA 1500
DB	682 CAAGGCAAAATGTTGGAGAGCAAGCAATTAATTTTACTTGGAGCAGATGTTTGGTAAACAGCC 741	1702 GTCATCATGACCAAAATGAGAACATTTACAGTGGTCAAGATAGGCTGCGAAGAGAGACCA 1761
QY	481 TGTAAATGTGCAATCAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAACAAT 540	1501 GAAACACACTCAACCAATCAGAGACAGAACTTAAACATTCATCTAACACATACCTTGCCT 1560
DB	742 TGTAAATGTGCAATCAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAACAAT 801	1762 GAAACACACTCAACCAATCAGAGACAGAACTTAAACATTCATCTAACACATACCTTGCCT 1821
QY	541 GAGTGGGTTACTTACTTACTTATAGCATAAGCAACCAACCAAGCACTATACCAAGTGT 600	1561 GGTGACTTTGTTGTTGTTGACAAAAATGAATGAGCCCATGATTTTGTGCTTGGTGGTAAAG 1620
DB	802 GAGTGGGTTACTTACTTACTTATAGCATAAGCAACCAACCAAGCACTATACCAAGTGT 861	1822 GGTGACTTTGTTGTTGTTGACAAAAATGAATGAGCCCATGATTTTGTGCTTGGTGGTAAAG 1881
QY	601 GTTCTTTTGGAAACATGTTGTTTACTTATTTTAACTAAAAAGAAATTAAGCACTAGT 660	1621 AATGTTTACCAATCTACCATGCAAGCTACTGTCTTAAATGAGGCGCAAGTTTCTGTATTGG 1680
DB	862 GTTCTTTTGGAAACATGTTGTTTACTTATTTTAACTAAAAAGAAATTAAGCACTAGT 921	1882 AATGTTTACCAATCTACCATGCAAGCTACTGTCTTAAATGAGGCGCAAGTTTCTGTATTGG 1941
QY	661 CCACCAAGAGACGGAGCTATTTTCTTAGCAGTACTCTGGCTGGAAAACTAATCTTTTAA 720	1681 TCAGAAAACTGGGCGGAGCCAAAGGTGCCAACTCTCTATAAATTTACTAGGTTGCGGACGC 1740
DB	922 CCACCAAGAGACGGAGCTATTTTCTTAGCAGTACTCTGGCTGGAAAACTAATCTTTTAA 981	1942 TCAGAAAACTGGGCGGAGCCAAAGGTGCCAACTCTCTATAAATTTACTAGGTTGCGGACGC 2001
QY	721 AAGAGAGGCGAGGCGCATCTAGTGAGCAAACTATACATGATGACATGCGGCGAGAAACG 780	1741 TCACCAATTCAGACACCGAAAAAGTACGCTCTCAGCCAGAACTATGCACTTAATCTCACTT 1800
DB	982 AAGAGAGGCGGAAACCGCATCTAGTGAGCAAAATTAACATGATGACATGCGGCGAGAAACG 1041	2002 TCACCAATTCAGACACCGAAAAAGTACGCTCTCAGCCAGAACTATGCACTTAATCTCACTT 2061
QY	781 GTTGAACCACTAATCACTGCGAGGAACTAAGCGGCGAGAAATTCAACTTAAATAA 840	1801 GCATCGGATCTCGAGGACCTGCTTTAGAGCTTTAGAGCCATGAGCAACCAATATCTCTGTTGCG 1860
DB	1042 GTTGAACCACTAATCACTGCGAGGAACTAAGCGGCGAGAAATTCAACTTAAATAA 1101	2062 GCATCGGATCTCGAGGACCTGCTTTAGAGCTTTAGAGCCATGAGCAACCAATATCTCTGTTGCG 2121
QY	841 GAAATGCTATTAAAACTACATTTAAAGAGCTGGTGATAAAAAGAGTAACCTCACAGAG 900	1861 GGCATCTGAGAAACCCAGAAACACTGCGGAGCTGTTTCCAAAGCTGCCAAGATGTTGTCAA 1920
DB	1102 GAGTTTCTATTAAAACTACATTTAAAGAGCTGGTGATAAAAAGAGTAACCTCACAGAA 1161	2122 GGCATCTGAGAAACCCAGAAACACTGCGGAGCTGTTTCCAAAGCTGCCAAGATGTTGTCAA 2181
QY	901 GACTGATGATGACGAGCAGACAGTTACATTTGAAATGATGCTCAACAGCTGGAGAA 960	1921 CTGAGCCCCAATTTGTTGAGAGGAGGATTTGAGAGCGTGTCTTCGTTGCGGAACCG 1980
DB	1162 GACTGATGATGACGAGCAGACAGTTACATTTGAAATGATGCTCAACAGCTGGAGAA 1221	2182 CTGAGCCCCAATTTGTTGAGAGGAGGATTTGAGAGCGTGTCTTCGTTGCGGAACCG 2241
QY	961 AACCTGCTGAAAAATAGCTTAGAGATTTGTACACTAATCTAGCCAGAACCAACCAAGCA 1020	1981 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019
DB	1222 AACCTGCTGAAAAATAGCTTAGAGATTTGTACACTAATCTAGCCAGAACCAACCAAGCA 1281	2242 TTGAAGAGAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2280

RESULT 9

MOU34253

LOCUS

MOU34253

4764 bp

DNA

linear

VRL 21-AUG-1996

[illegible]

Db 1802 ACAGAAAACCTGGGGGACCGAAGTGGCCGACTCCTATATAAATCACTAGTTCGGCAGCG 1861

Qy 1741 TCACCAATTCAGACACCGAAAAGTACGCTCTCAGCGACCAACTATGCACCTAACCTCACTT 1800

Db 1862 TCACCAATTCAGACACCGAAAAGTACGCTCTCAGCGACCAACTATGCACCTAACCTCACTT 1921

Qy 1801 GCATCGGATCTCAGGACCTGGCTTTAGAGCCTTGGAGCACCACCAATATCTCTGTTGG 1860

Db 1922 GCATCGGATCTCAGGACCTGGCTTTAGAGCCTTGGAGCACCACCAATATCTCTGTTGG 1981

Qy 1861 GGCACTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGGTCAA 1920

Db 1982 GGCACTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGGTCAA 2041

Qy 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGACGCTGCTTCGGTGGGAAACCG 1980

Db 2042 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGACGCTGCTTCGGTGGGAAACCG 2101

Qy 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACCTTGACTAA 2019

Db 2102 TTGAAGAAAGACTTCAGCGAGCGCTGAACCTTGACTAA 2140

RESULT 10

PAMVMI

LOCUS

DEFINITION Mouse parvovirus minute virus immunosuppressive variant genome (= MVMi).

ACCESSION X02481

VERSION GI:60918

KEYWORDS coat protein; genome; origin of replication; overlapping genes; terminal repeat; unidentified reading frame.

SOURCE Mouse minute virus

ORGANISM Mouse minute virus

REFERENCE 1 (bases 1 to 5087)

AUTHORS Sahli,R., McMaster,G.K. and Hirt,B.

TITLE DNA sequence comparison between two tissue-specific variants of the autonomous parvovirus, minute virus of mice

JOURNAL Nucleic Acids Res. 13 (10), 3617-3633 (1985)

MEDLINE 85242059

PUBMED 3855242

COMMENT For the fibroblast-specific strain (MVMp) sequence see <PAMW2>. The genomes of MVMp and MVMi (immunosuppressive variant) have more than 96% of their sequence in common.

DATA kindly reviewed (18-JUL-1986) by G. McMaster.

FEATURES

source 1..5087

organism="Mouse minute virus"

mol_type="genomic DNA"

db_xref="taxon:10794"

misc_feature 1..116

notes="terminal hairpin region"

promoter 176..183

notes="TATA box"

gene 2405..4558

genes="VP1"

CDS <2405..4558

gene="VP1"

functions="coat protein"

codon_start=1

product="VP1 protein"

protein_id="CAB46507.1"

db_xref="GI:5419928"

db_xref="GOA:P07302"

translation="VPPGYKLGNSLDQGEPTNPSDAAKSHDBAYDQYIKSGKNP YLYFSADQRFIDQTKADKGVGHYFFRTKRAFAPKLATDSBEFTSGVSRAGRT RPAYTIFINARAKKLTSSAAQSSQSTMSDGTSPQDGGNGVHSAARAAADPGGS GGGGGGGVSTGVSVDNCTHYRFLGDGWEITATATLRLVHLMNPKESEYCRIRVH TTDTSVKGNWAKDDAHEQIWPMSLVANAGVWLQPSDWQYICNTMSQLNLSLQOE IFNVLTVEQDSGQAIKIYNDLITACMVAVDNSNLLPYTPAANSMETLIGFYFWK PTIASPYRYFCVDRDLSVYENQEGTIEHNVMGTPKGMNSQFFTIENQIITLLRTG

DEPATCTYYFDNPNVKLTHTWQTNROLGPPILLSTPEADTDAAGTLTAQSSRHGATQM EVNWSEAIRTPAQVGFQPHNDPEASRAGPFAAPKVPADVTQGVREANGSVYSY KGDNENWAAHGPAPEPTWDETDFGSDTRDGFIOAPLVVPPPLNGILTNWNPAG TNDIHFNSVNSYGLPTAFSHSPSPYVQGOIWDKELDLEHKLPHLITAFVCKNAP GQMLVRLGNLTDQVDPNGATLSRIYVYGTFFWKGLTMRKLRANTTNWNPVQVSVE DNGNSVMSVTKWLPTATGNMQSVPLITRPARNTY"

2792..4558

gene

CDS <2792..4558

gene="VP2"

functions="coat protein"

codon_start=1

product="VP2 protein"

protein_id="CAB46508.1"

db_xref="GI:5419929"

db_xref="GOA:P07302"

translation="TMSDGTSPDGGNGVHSAARVRAADPGSGGGGGGGVGVSV TGSYDNQTHYRFLGDGWEITATATLRLVHLMNPKESEYCRIRVHNTTDTSVKGNMAK DAHEQIWPMSLVANAGVWLQPSDWQYICNTMSQLNLSLQOEIPNVVLKVTBQD SGQAIKIYNDLITACMVAVDNSNLLPYTPAANSMETLIGFYFWKPTIASPYRYFCV DRDLSVYENQEGTIEHNVMGTPKGMNSQFFTIENQIITLLRTGDEFAFGTYFPTN PVKLTHTWQTNROLGPPILLSTPEADTDAAGTLTAQSSRHGATQMEVNWSEAIRTP AOVGFCOPHNDPEASRAGPFAAPKVPADVTQGVREANGSVYSYKGHGENWAAHGP APERTWDETDFGSDTRDGFIOAPLVVPPPLNGILTNWNPICGTKNDIHFNSVNS YGLPTAFSHSPSPYVQGOIWDKELDLEHKLPHLITAFVCKNAPGQMLVRLGNLTD QYDNGATLSRIYVYGTFFWKGLTMRKLRANTTNWNPVQVSVEDNGNSYMSVTKWL PTATGNMQSVPLITRPARNTY"

4603..4608

misc_feature

note="polyadenylation signal"

4821..4826

misc_feature

note="polyadenylation signal"

4878..5087

misc_feature

note="terminal hairpin region"

ORIGIN

Query Match 94.2%; Score 1902.2; DB 14; Length 5087;

Best Local Similarity 96.4%; Pred. No. 0;

Matches 1946; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

Qy 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTCGGGACCAACCACTGGTTAAAGGAAAA 60

Db 262 ATGGCTGGAAATGCTTACTCTGATGAAGTTCGGGACCAACCACTGGTTAAAGGAAAA 321

Qy 61 AGTAACCAAGAGTGTCTCATTTGTTTAAATAATGAAATGTTCAACTGAATGGAATA 120

Db 322 AGTAACCAAGAGTGTCTCATTTGTTTAAATAATGAGGATGTTCAACTGAATGGAATA 381

Qy 121 GATATCGGATGGAATAGTTACAAAAAGAGCTGCAGAGACGAGCTGAAATCTTTTCAA 180

Db 382 GATATCGGATGGAATAGTTACAAAAAGAGCTGCAGAGACGAGCTGAAATCTTTTCAA 441

Qy 181 CGAGGAGCGAAACTACTTGGGACCAAGAGCTGCAGAGACGAGCTGAAATCTTTTCAA 240

Db 442 CGAGGAGCGAAACTACTTGGGACCAAGAGCTGCAGAGACGAGCTGAAATCTTTTCAA 501

Qy 241 GAAATGACCAAAAAAGCAAGTATTTCATTTTGGTTTCTTTGTTTAAAAAATGTTTATTTGAA 300

Db 502 GAAATGACCAAAAAAGCAAGTATTTCATTTTGGTTTCTTTTGGTTTAAAAAATGTTTATTTGAA 561

Qy 301 GTGCTTAAACAAAAAGATATATTTTCCTGCTGATGTTTAAATGTTGTTGCAACATGAATGG 360

Db 562 GTGCTTAAACAAAAAGATATATTTTCCTGCTGATGTTTAAATGTTGTTGCAACATGAATGG 621

Qy 361 GGAAGAGACCAAGCTGGCACTGCCATGTAATAATTTGGAGGAAGGACTTTTAGTCAAGCT 420

Db 622 GGAAGAGACCAAGCTGGCACTGCCATGTAATAATTTGGAGGAAGGACTTTTAGTCAAGCT 681

Qy 421 CAAGGAAATGGTGGGAAAGGCAACTAAATGTTTACTTGGAGCAGATGTTTGGTAAACAGCC 480

Db 682 CAAGGAAATGGTGGGAAAGGCAACTAAATGTTTACTTGGAGCAGATGTTTGGTAAACAGCC 741

Qy 481 TGTAATGTGCAACTAACACAGCTGAAAGAAATTAAGAAATAGAGAAATAGCAGAGACAAT 540

OTKKEYSIKTTIKELVHKRVTSBEDWMMOPDSYIEWMAQPGGNNLKNLTLEICTLTL
ARTYFDLLEKAEKTSKLTNFSLPDTRICKYIFAPHGWNITKVCCHCAVCLNKKQGGKR
NTVLFHPSGTGSIITIAQAVGNVCKYIAANVFPNPFNDCKNKLWYEEAGNFGQ
QNOFAIKCSGTIRIDQKGGKQIEPTVIMTNTNENITVVKIKGGEERPEHTQPIRD
RMLNHLTHLPDGFGLVDKNWPMICAWLVKNGYOSWASVCAKRGKVPDWTENWAE
PKVPTPINSIGSARSPTPKPSPLSONVALTPLASDLLEDLALPEKSTNTVPVGTAE
TQNTGAGSKACODGQLSPWBSIEBDRALCTFQAEPLKXDFSEPLND"
join(2290..2320,2403..4561)
/gene="vpl1"
CDS
join(2290..2320,2403..4561)
/gene="vpl1"
/codon_start=1
/product="capsid protein"
/protein_id="AAA61406.1"
/db_xref="GI:525327"

/translations="MAPPKAKRKRKRWVPGYKYLPGCNLSLDQGEPTNPSPDAAKEHDE
AYDKYIKSGNPLVYFSADOREIDOTKADWGGKVGHYFFRTKFAFAPRLASSSEP
GTSVSIAGKRTKPPAHIFINOARAKKRLAAQOORTOTMSDAGOPSGSVAQSA
RVBRADPGSGSGSGGGGVSTGSVDNTHYRFLSDGWVEITAYSTRVHLMRP
KSENYCRVRHNTNRTAGNMAKDDAHEQIWPNSLIDSNANGVWFQPSDWOFICNN
MSHVNLSLDQELFNVIKTVTEONTGAEXAVKIYNNDLTASMMVALDSNNILPYTPAT
DNQETLGFYPMKFTMPSVRYFYFNCDSLSVTYDQTSIVDTMANASGLSOFPIE
NTORIOLLRTGDSFATGTYFETPEPIKLSHTWOSNRQLGPOPOITDLPADNENATLV
TRGDRSGIOTISGNSNVTBATRVRPAQVGFQPHDNFETSRAGFPKVPVPAVNTQGN
EHDANGSLRYTDKQGDWGSNNKERTWDALSYDSGRWADRCFINATPPTSPPAL
NNLTNDSPIGNKTAIHYQNVFNSYGLPTAFPHPAPIYPOGQIWDKELDLHKKPLRHA
QAPFVCKNAPGQLLALPNLTDQYDPSNSTLSRYTYGTFPFWKGLTKAKLRPNA
TWNPVYQVSAQYQNEVENS IHKWLPTATGNMQSIPLLSRPVARNTY"

ORIGIN

Query Match	94.1%;	Score 1899;	DB 14;	Length 5144;
Best Local Similarity	96.3%;	Pred. No. 0;		
Matches 1944;	Conservative 0;	Mismatches 75;	Indels 0;	Gaps 0;
QY	1	ATGGCTGGAAATGCTTACTCTGATGAAGTTTTCGGAGCAACCAACTGGTTTAAAGGAAAA	60	
DB	265	ATGGCTGGAAATGCTTACTCTGATGAAGTTTTCGGAGCAACCAACTGGTTTAAAGGAAAA	324	
QY	61	AGTAACAGGAGTGTCTCATTTGTTTTTAAAAATGTTCAACTGAAATGGAAAA	120	
DB	325	AGTAACAGGAGTGTCTCATTTGTTTTTAAAAATGTTCAACTGAAATGGAAAA	384	
QY	121	GATATCGGATGGAATAGTTACAAAAGAGCTCGAGGAGCAGCTGAAATCTTTACAA	180	
DB	385	GATATCGGATGGAATATTAACAAAAGAGCTCGAGGAGCAGCTGAAATCTTTACAA	444	
QY	181	CGAGGAGCGAACTACTTCGGACCAAGAGGAGCATGGAATGGGAAACACAGTGGAT	240	
DB	445	CGAGGAGCGGAACTACTTCGGACCAAGAGGAGCATGGAATGGGAACTACAGTGGAT	504	
QY	241	GAAATGACAAAAGCAAGTATTCAATTTGATTCTTTGGTTAAAAAATGTTTATTGAA	300	
DB	505	GAAATGACAAAAGCAAGTATTCAATTTGATTCTTTAGTTAAAAAATGTTTATTGAA	564	
QY	301	GTGCTTAAACAAAAGAAATATTTCTGTTGATGTTTAATTTGGTTTGGCAACATGAATGG	360	
DB	565	GTGCTTAGCAACAAAGAAATATAGCTCTCTGCTGATGTTACTTTGGTTCGAGCATGAATGG	624	
QY	361	GGAAAAGACCAAGCTGGCACTGCCATGTACTTAATTTGGAGGAAGGACTTTTGTCAAGCT	420	
DB	625	GGAAAAGACCAAGCTGGCACTGCCATGTACTTAATTTGGAGGAAGGACTTTTGTCAAGCT	684	
QY	421	CAAGGGAATATGTTGGAGAGGCAACTAAATATGTTTACTTGGAGCAGATGTTGGTAAACAGCC	480	
DB	685	CAAGGGAATATGTTGGAGAGGCACTAAATATGTTTACTTGGAGCAGATGTTGGTAAACAGCC	744	
QY	481	TGTAATGTGCAACTTAACACCAAGCTGAAAGAAATTAACCTAAAGAAATAGCAGAGCAAT	540	
DB	745	TGTAATGTGCACTTAACACCAAGCTGAAAGAAATTAACCTAAAGAAATAGCAGAGCAAT	804	
QY	541	GAGTGGGTTACTTACTTACTTATAGCATAAGCAACCAAAAGCACTATACCAAGTGT	600	
DB	805	GAGTGGGTTACTTACTTACTTATAGCATAAGCAACCAAAAGCACTATACCAAGTGT	864	

QY	601	GTTCTTTTGGAAAAATGATTTGCTTACTATATTTTAACTAAAAAGAAAAATAAGCACTAGT	660	
DB	865	GTTCTTTTGGAAAAATGATTTGCTTACTATATTTTAACTAAAAAGAAAAATAAGCACTAGT	924	
QY	661	CCACCAAGAGACGGAGGCTATTTTCTTAGCAGTGACTCTCGCTGGAAAACTAACTTTTAA	720	
DB	925	CCCCCAAGAGACGGAGGCTATTTTCTTAGCAGTGACTCTCGCTGGAAAACTAACTTTTAA	984	
QY	721	AAAGAGGGGAGCGCCATCTAGTGAGCAAACTATACATCTGATGACATCGCGCCAGAAACG	780	
DB	985	AAAGAGGGGAGCGCCATCTAGTGAGCAAACTATACATCTGATGACATCGCGCCAGAAACG	1044	
QY	781	GTTGAAACCAACAGTAACCACTGCGCAGGAAAACCTAAGCGCGGAGAAATTTCAAACTAAAA	840	
DB	1045	GTTGAAACCAACAGTAACCACTGCGCAGGAAAACCTAAGCGCGGAGAAATTTCAAACTAAAA	1104	
QY	841	GAAGTTGCTATTAAAACTACACTTAAAGAGCTGCTGTCATAAAAAGAGTAACCTCACACAG	900	
DB	1105	GAGGTTTCTATTAAAAACCACTTAAAGAGCTGCTGTCATAAAAAGAGTAACCTCACACAG	1164	
QY	901	GACTGGATGATGATGCGAGCCAGACAGTTACATTTGAAATGATGGCTCAACAGGTGGAGAA	960	
DB	1165	GACTGGATGATGATGCGAGCCAGACAGTTACATTTGAAATGATGGCTCAACAGGTGGAGAA	1224	
QY	961	AACCTGCTGAAAAATACGGCTAGAGATTGTACACTAATCTTAGCCAGAACCAAAACAGCA	1020	
DB	1225	AACCTGCTGAAAAATACGGCTAGAGATTGTACACTAATCTTAGCCAGAACCAAAACAGCA	1284	
QY	1021	TTTGTACTTAAATTTAGAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCTGAC	1080	
DB	1285	TTTGTACTTAAATTTAGAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCTGAC	1344	
QY	1081	ACAAGAACCTGCAAGAAATTTTGTCTTTTTCATGGCTGGAATCTATGTTAAAGTTTGCATGCT	1140	
DB	1345	ACAAGAACCTGCAAGAAATTTTGTCTTTTTCATGGCTGGAATCTATGTTAAAGTTTGCATGCT	1404	
QY	1141	ATTGCTGTGTTTTTAAACAGACAAAGGAGGCAAAAGAAATACTGTTTATTTTATTCATGACCA	1200	
DB	1405	ATTGCTGTGTTTTTAAACAGACAAAGGAGGCAAAAGAAATACTGTTTATTTTATTCATGACCA	1464	
QY	1201	GCAGACACAGGCAAAATCTATTATTGACACAGCCATAGCAACAGCTTGGCAATGTTGGT	1260	
DB	1465	GCAGACACAGGCAAAATCTATTATTGACACAGCCATAGCAACAGCTTGGCAATGTTGGT	1524	
QY	1261	TGCTATAATGACGCAATGTAAACTTTTCAATTTTAACTGTATACCAACAGAACTTGAT	1320	
DB	1525	TGCTATAATGACGCAATGTGAATTTTCCATTTCAATGACTGTATACCAACAGAACTTAAT	1584	
QY	1321	TGGGTAGAGAAAGCTGGTAACTTTTGGACACAGCAAGTAAACCCAGTTTAAAGCCATTGCTCT	1380	
DB	1585	TGGGTAGAGAAAGCTGGTAACTTTTGGACACAGCAAGTAAACCCAGTTTAAAGCCATTGCTCT	1644	
QY	1381	GGTCAAACTATTTCGCACTGATCAAAAGGAAAGGAGGCAAAACAGATTGAAACCAACCA	1440	
DB	1645	GGTCAAACTATTTCGCACTGATCAAAAGGAAAGGAGGCAAAACAGATTGAAACCAACCA	1704	
QY	1441	GTCAATCATGACCAAAATGAGAACTTACAGTGGTCAGAAATAGGCTGCGAGAAAGACCA	1500	
DB	1705	GTCAATCATGACCAAAATGAGAACTTACAGTGGTCAGAAATAGGCTGCGAGAAAGACCA	1764	
QY	1501	GAAACACTCAACCAATCAGACAGAAATGCTTAAACATTTCATCTAACACATACATCTGCCT	1560	
DB	1765	GAAACACTCAACCAATCAGACAGAAATGCTTAAACATTTCATCTAACACATACATCTGCCT	1824	
QY	1561	GGTGACTTTGGTTGGTTGACAAAATGATGGCCCATGATTTGCTGCTGGTTGGTAAAG	1620	
DB	1825	GGTGACTTTGGTTGGTTGACAAAATGATGGCCCATGATTTGCTGCTGGTTGGTAAAG	1884	
QY	1621	AATGGTTTACCAATCTACCACTAGCAAGCTACTGTGCTTAAATGGGGCAAAAGTTCTGATTGG	1680	
DB	1885	AATGGTTTACCAATCTACCACTAGCAAGCTACTGTGCTTAAATGGGGCAAAAGTTCTGATTGG	1944	

```
QY 1681 TCAGAAACTGGCGGAGCCAAAGGTGCCAACTCCTATAAATTTATAGTTCGGCAGCG 1740
Db 1945 ACGAAAACCTGGCGGAGCCGAAGGTGCCACTCTATAAATTTCACTAGGTTCGGCAGCG 2004
QY 1741 TCACATTCACGACACCCAGAAAGTACGCTCTCAGCCAGAACATATGCACTAACTCCACTT 1800
Db 2005 TCACATTCACGACACCCAGAAAGTACGCTCTCAGCCAGAACATATGCACTAACTCCACTT 2064
QY 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGCAGACACCAATATCTCTGTTGCG 1860
Db 2065 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGCAGACACCAATATCTCTGTTGCG 2124
QY 1861 GGCATCTCAGAAACCCAGAACACACTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGFTCAA 1920
Db 2125 GGCATCTCAGAAACCCAGAACACACTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGFTCAA 2184
QY 1921 CTGAGCCCAACTGGTCAGAGATCGAGGAGGATTTGAGACGCTGCTTCGGTGGGAACCG 1980
Db 2185 CTGAGCCCAACTGGTCAGAGATCGAGGAGGATTTGAGACGCTGCTTCGGTGGGAACCG 2244
QY 1981 TTGAAGAAAGACTTCAGCGAGCGCGCTGAACTTGGACTAA 2019
Db 2245 TTGAAGAAAGACTTCAGCGAGCGCGCTGAACTTGGACTAA 2283

RESULT 12
MOU34254
LOCUS Mouse parvovirus 1c DNA. 4764 bp DNA linear VRL 21-AUG-1996
ACCESSION U34254
VERSION U34254.1 GI:1464794
KEYWORDS Mouse parvovirus 1c
SOURCE Mouse parvovirus 1c
ORGANISM Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.

REFERENCE 1 (bases 1 to 4764)
AUTHORS Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L.,
Franklin,C.H., Hook,R.R. Jr. and Riley,L.K.
TITLE Molecular characterization of newly recognized rodent parvoviruses
JOURNAL J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
MEDLINE 96201434
PUBMED 8609486

REFERENCE 2 (bases 1 to 4764)
AUTHORS Besselsen,D.G.
TITLE Direct Submission
JOURNAL Submitted (17-AUG-1995) David G. Besselsen, Department of
Veterinary Pathology, University of Missouri-Columbia, W213
Veterinary Medicine Building, Columbia, MO 65211, USA

FEATURES
source
1. .4764
/organism="Mouse parvovirus 1c"
/mol_type="genomic DNA"
/specific_host="Mus musculus"
/db_xref="taxon:42842"

ORIGIN
Query Match 92.8%; Score 1873; DB 14; Length 4764;
Best Local Similarity 95.4%; Pred. No. 0;
Matches 1927; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 60
Db 122 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAAACCAACCACTGGTTAAAGGAAAA 181
QY 61 AGTAACACGAGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGGAATGAAAA 120
Db 182 AGTAACACGAGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGGAATGAAAA 241
QY 121 GATATCGGATGGAATGTTTCAAAAAAGAGCTCGAGGAGGACGAGCTGAAATCTTTTACAA 180
Db 242 GATATTTGGATGGAATGTTTCAAAAAAGAGCTCGAGGAGGACGAGCTGAAATCTTTTACAA 301
QY 181 CGAGGAGCGGAACTACTTGGGACCAAGCGGAGGACATGGAATGGGAAACCACTGGAT 240
```

```
Db 302 CGAGGAGCGGAACTACTTGGGACCAAGCGGAGGACATGGAATCTACAGTGGAT 361
QY 241 GAAATCACCACCAAAAGCAAGTATTTCATTTTGGTTGTTTAAAAAATGTTTATTGAA 300
Db 362 GAAGTACACCAAAAGCAAGTATTTCATTTAGACTCTTTAGTTAAAAAATGTTTATTGAA 421
QY 301 GTGCTTAAACACAAAGAAATATATTTCTCTGGTGTATGTTAAATTTGGTTTGTGCAACATGAATGG 360
Db 422 GTGCTTAAACACAAAGAAATATATTTCTCTGGTGTATGTTAAATTTGGTTTGTGCAACATGAATGG 481
QY 361 GGAAGAAGCAAGGCTGGCACTGCCATGTACTAAATTTGGAGGAAAGCACTTTAGTCAAGCT 420
Db 482 GGAAGAAGCAAGGCTGGCACTGCCATGTACTAAATTTGGAGGAAAGCACTTTAGTCAAGCT 541
QY 421 CAAGGCAAAATGGTGGAGAAAGCAACTAAATTTTACTGTGAGCAGATGTTGGTAAACAGCC 480
Db 542 CAAGGCAAAATGGTGGAGAAAGCACTAAATTTTACTGTGAGCAGATGTTGGTAAACAGCC 601
QY 481 TGTAAATGTGCAACTAACACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAAGCAAT 540
Db 602 TGTAAATGTGCAACTAACACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAAGCAAT 661
QY 541 GAGTGGGTACTCTACTTACTTAAAGCATAAGCAAAACCAAAAGCACTATACCAAGTGT 600
Db 662 GAATGGGTTTACCTTGCTCACTTATAAGCATAAGCAAAACCAAAAGCACTATACCAAGTGT 721
QY 601 GTTCTTTTGGAAAAACATGATGCTTACTTACTTATTTTAACTTAAAGAAAAATAGCACTAGT 660
Db 722 GTTCTTTTGGAAAAACATGATGCTTACTTACTTATTTTAACTTAAAGAAAAATAGCACTAGT 781
QY 661 CCACCAAGAGACGAGGCTTATTTCTTAGCAGTGAATCTGGCTGGAAAACTAACTTTTAA 720
Db 782 CCGCCAAAGGACGAGGCTTATTTCTTAGTGTACTCTGGCTGGAAAACTAACTTTTAA 841
QY 721 AAAGAAGGCGAGCGGCTCTAGTGACAACTATACACTGATGACATGCGGCGCAGAAACG 780
Db 842 AAAGAAGGCGAGCGGCTCTAGTGACAACTATACACTGATGACATGCGGCGCAGAAACG 901
QY 781 GTTGAACACCACTAGTAACTCTCGCAGGAAACTAAGCGCGGAGAAATTCAAACTTAAAAA 840
Db 902 GTTGAACACCACTAGTAACTCTCGCAGGAACTAAGCGCGGAGAAATTCAAACTTAAAAA 961
QY 841 GAAGTGTCTTATTAACCTACCTTAAAGAGCTGGTGCAATAAAGAGTAACTCACACAGAG 900
Db 962 GAGGTTTCTATTATAAACACACACTTAAAGAGCTAGTGCAATAAGAGAGTAACTCACACAGAA 1021
QY 901 GACTGGATGATGATGCGCAGCAGACAGTTACATTTGAAATGATGCTCAACACAGGTGGAGAA 960
Db 1022 GACTGGATGATGATGCGCAGCAGACAGTTATATTTGAAATGATGCTCAACACAGGTGGAGAA 1081
QY 961 AACCTGTGAAAAATACCGCTAGAGATTTGTACACTAACTCTAGCCAGAACCAAAACAGCA 1020
Db 1082 AACCTGTGAAAAATACCGCTAGAGATCTGTACACTTAACCTCTAGCTAGAACCAAAACAGCA 1141
QY 1021 TTTGACTTAAATTTAGAAAAAGCTGAAACACAGCAAACTAACCAACTTTTCACTGCTGAC 1080
Db 1142 TTTGACTTAAATTTAGAAAAAGCTGAAACACAGCAAACTAACCAACTTTTCACTGCTGAC 1201
QY 1081 ACAAGAACTGAGAAATTTTGTCTTCTGCTGGAATATGTTAAAGTTTGCATGCT 1140
Db 1202 ACAAGAACTGAGAAATTTTGTCTTCTGCTGGAATATGTTAAAGTTTGCATGCT 1261
QY 1141 ATTTGCTGTGTTTAAACAGACAGGAGCAAAAGAAATACCTGTTTATTTTTCATGGACCA 1200
Db 1262 ATTTGCTGTGTTTAAACAGACAGGAGCAAAAGAAATACCTGTTTATTTTTCATGGACCA 1321
QY 1201 GCCAGCAGCGCAATCTATTATTCACAGCCATGACAGCAAGAGCTGGCAATGTGTG 1260
Db 1322 GCCAGTACAGGCAATCCATTTATTGCAACAGCCATGACAGCGAGCTGGTAAATGTGTG 1381
QY 1261 TGCCTAATGACCAATGTTAAACTTCCATTTAATGACTGTACCAACAGAGACTGATT 1320
```

Db 1382 TGTATAATGACGAAATGTGAATCTTCCATTCAATGACTGTACCAACGAAGACTTAATT 1441
Qy 1321 TGGGTAGAAGAGCTGTAACTTTGGACAGCAAGTAACACAGTTTAAAGCCATTGTCTCT 1380
Db 1442 TGGGTGGAAGAGCTGTAACTTTGGACAGCAAGTAACACAGTTTAAAGCCATTGTCTCT 1501
Qy 1381 GGTCAAACTATTCCGATTTGATCAAAAGGAAAGGAGCAAGCAAGATTTGAACCAACACCA 1440
Db 1502 GGTCAAACTATTCCGATTTGATCAAAAGGAAAGGAGCAAGCAAGATTTGAACCAACACCA 1561
Qy 1441 GTCATCATGACCAAAATGAGAACATTACAGTGTGTCAGATAGGCTGCGAAGAAAGACCA 1500
Db 1562 GTCATCATGACCAAAATGAGAACATTACAGTGTGTCAGATAGGCTGCGAAGAAAGACCA 1621
Qy 1501 GAACACACTCAACCAATCAGACAGACAGATGCTTAAACATTCACTTAACACATACATTGCCT 1560
Db 1622 GAACACACTCAACCAATCAGACAGACAGATGCTTAAACATTCACTTAACACATACATTGCCT 1681
Qy 1561 GGTGACTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGTGCTTGGTTGGTAAAG 1620
Db 1682 GGTGACTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGTGCTTGGTTGGTAAAG 1741
Qy 1621 AATGGTTTACCAATCTACATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTCTGATTGG 1680
Db 1742 AATGGTTTACCAATCTACATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTCTGATTGG 1801
Qy 1681 TCAGAAAATCTGGCGGAGCCAAAGTGCCTTCTTAAATTTACTAGTTTCGGACGCG 1740
Db 1802 ACGGAAAATCTGGCGGAGCCAAAGTGCCTTCTTAAATTTCACTAGTTTCGGACGCG 1861
Qy 1741 TCACCAATTCAGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTTAACCTCACTT 1800
Db 1862 TCACCAATTCAGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTTAACCTCACTT 1921
Qy 1801 GCATCGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGACACACCAATATCTCTGTGTCG 1860
Db 1922 GCATCGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGACACACCAATATCTCTGTGTCG 1981
Qy 1861 GGCACCTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGTTCAA 1920
Db 1982 GGCACCTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGTTCAA 2041
Qy 1921 CTGAGCCCAACTGGTCAGAGATCGAGAGAGATTTGAGACGCTGCTTCGTTGGGAACCG 1980
Db 2042 CTGAGCCCAACTGGTCAGAGATCGAGAGAGATTTGAGACGCTGCTTCGTTGGGAACCG 2101
Qy 1981 TTGAAGAGACTTCAGCGAGCCGCTGAACCTGACTAA 2019
Db 2102 TTGAAGAGACTTCAGCGAGCCGCTGAACCTGACTAA 2140

RESULT 13

LOCUS HOU34255 4773 bp DNA linear VRL 21-AUG-1996
DEFINITION Hamster parvovirus DNA.
ACCESSION U34255
VERSION U34255.1 GI:1464792
KEYWORDS
SOURCE Hamster parvovirus
ORGANISM Hamster parvovirus
REFERENCE 1 (bases 1 to 4773)
Besselsen, D.G., Pintel, D.J., Purdy, G.A., Besch-Williford, C.L.,
Franklin, C.L., Hook, R.R. Jr. and Riley, L.K.
Molecular characterization of newly recognized rodent parvoviruses
J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
MEDLINE 96201434
PUBMED 8609486
REFERENCE 2 (bases 1 to 4773)
Besselsen, D.G.
Direct Submission
TITLE Submitted (17-AUG-1995) David G. Besselsen, Department of
Veterinary Pathology, University of Missouri - Columbia, W213

Veterinary Medicine Building, Columbia, MO 65211, USA

FEATURES
source

1. 4773
/organism="Hamster parvovirus"
/mol_type="genomic DNA"
/specific_host="Mesocricetus auratus"
/db_xref="taxon:42843"

ORIGIN

Query Match 92.2%; Score 1862.2; DB 14; Length 4773;

Best Local Similarity 95.1%; Pred. No. 0;

Matches 1921; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

Qy 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTTCGGAGCAACCACTGGTTTAAAGGAAAAA 60

Db 122 ATGGCTGGAAATGCTTACTCTGATGAAGTTTTCGGAAACAACCACTGGTTTAAAGGAAAA 181

Qy 61 AGTAACACGAGGAAGTGTCTCATTTTGTAAAAATGMAAATGTTCAACTGAATGGAAAA 120

Db 182 AGTAACACGAGGAAGTGTCTCATTTTGTAAAAATGMAAATGTTCAACTGAATGGAAAA 241

Qy 121 GATATCGGATGGAAATAGTTTACAAAAAGAGCTCCAGGAGGACGAGCTGAAATCTTTACAA 180

Db 242 GATATCGGATGGAAATAGTTTACAAAAAGAGCTCCAGGAGGAGAGCTGAAATCTTTACAA 301

Qy 181 CGAGGAGCGGAACTACTTGGGACCAAGCGGAGGACATGGAATGGGAAACCAACAGTGGAT 240

Db 302 CGAGGAGCGGAACTACTTGGGACCAAGCGGAGGAGATGGAATGGGAAATCTTCACTGGAT 361

Qy 241 GAAATGACCAAAAGCAAGTATTCTTTTGTGTTTAAATGTTTAAATGTTTAAATGTTTAA 300

Db 362 GAAATGACCAAAAGCAAGTATTCTTTTGTGTTTAAATGTTTAAATGTTTAAATGTTTAA 421

Qy 301 GTGCTTAAACAAGAAATATATTTCTCTGGTGTATGTTTAAATGTTTGTGCAACATGAATGG 360

Db 422 GTGCTGATACAAAGAACATAGCACCTAGTGAATGTTTGTGTTGTGATGATGATGATGAT 481

Qy 361 GGAAGACCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAAGAGCTTTAGTCAAGCT 420

Db 482 GGAAGACCAAGGCTGGCACTGTCTGATCTAATTTGGAGGCAAGGACTTTAGCCAAGCT 541

Qy 421 CAGGGAATGTTGGAGAGGCACTAATTTTCTGAGGAGGAGTGGTTGGTAAACAGCC 480

Db 542 CAAGGAAATGTTGGAGAGGCACTAATTTTCTGAGGAGGAGTGGTTGGTAAACAGCC 601

Qy 481 TGTAAATGTCAACTAAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT 540

Db 602 TGTAGTGTGAGCTATTAACAGCTGAAAGAAATTAAGCTGAGAGAGATAGCGGAGACCA 661

Qy 541 GAGTGGTTTACTCTTACTTATAGCAATAAGCAAAACCAAAAGACCTATATACCAAGTGT 600

Db 662 GAATGGTCACTTTGCTTACTTATAGCAATAAGCAAAACCAAAAGACCTATATACCAAGTGT 721

Qy 601 GTTCTTTTGGAAACATGATGCTTACTTATTTTAACTTAAAGAAATTAAGCACTAGT 660

Db 722 GTTGTCTTTGGAAATAGTGTGCTTACTTATTTTATCCAAAGAGAAATATGACCAGT 781

Qy 661 CCACCAAGAGAGCGAGGCTATTTCTTAGCAGTACTCTGCTGGGAAAACTAACTTTTAA 720

Db 782 CCACCAAGAGAGCGAGGCTATTTCTTAGCAGTACTCTGCTGGGAAAACTAACTTTTAA 841

Qy 721 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACATCTGATGACATGCGGCGAGAAACG 780

Db 842 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACATCTGATGACATGCGGCGAGAAACG 901

Qy 781 GTTGAACCAACAGTAACCACTGCGGAGGAACTAAGCGGCGAGATTCGAATCTAAATAA 840

Db 902 GTTGAACCAACAGTAACCACTGCGGAGGAACTAAGCGGCGAGATTCGAATCTAAATAA 961

Qy 841 GAAGTTGCTATTAAAACTACACTTAAAGAGCTGCTGCATAAAAAGAGTAACCTCACCAGAG 900

Db 962 GAGGTCTCTATTAAAAACCACTTAAAGAGCTGCTGCATTAAGAGATGATCTCACCAGAA 1021


```
QY 901 GACTGATGATGATGAGCAGCAGTACATTGAAATGATGCTCAACAGAGTGAGAA 960
Db 1022 GACTGATGATGATGAGCAGCAGTACATTGAAATGATGCTCAACAGAGTGAGAA 1081
QY 961 AACCTGCTGAAATAGCTAGAGATTGTACACTAACTCTAGCCAGAACCAACAGCA 1020
Db 1082 AACCTGCTGAAATAGCTAGAGATTGTACACTAACTCTAGCCAGAACCAACAGCA 1141
QY 1021 TTTGACTTAATTTAGAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCTGAC 1080
Db 1142 TTTGACTTAATTTAGAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCTGAC 1201
QY 1081 ACAAGAACTGCAAGATTTTGTCTTTTCATGCTGGAATATCTTAAAGTTTCCATGCT 1140
Db 1202 ACAAGAACTGCAAGATTTTGTCTTTTCATGCTGGAATATCTTAAAGTTTCCATGCT 1261
QY 1141 ATTTGCTGTGTTTAAACAGACAGAGGCAAGAAAGAAATCTGTTTATTTATTCATGACCA 1200
Db 1262 ATTTGCTGTGTTTAAACAGACAGAGGCAAGAAAGAAATCTGTTTATTTATTCATGACCA 1321
QY 1201 GCCAGCACAGGCAAACTATTATTGCAAGCCATAGCAACAGCTTGGCAATGTTGCT 1260
Db 1322 GCCAGTACAGCAAACTATTATTGCAAGCCATAGCAACAGCTTGGTAAATGTTGCT 1381
QY 1261 TGCTATAATGCAGCAATGTAAACTTTTCCATTTAATCACTGTACCAACAGAACTTGATT 1320
Db 1382 TGCTATAATGCAGCAATGTAAACTTTTCCATTTAATCACTGTACCAACAGAACTTGATT 1441
QY 1321 TGGGTAGAGAAGCTGGTAACTTTTGACAGCAAGTAAACAGTTTAAAGCCATTGCTCT 1380
Db 1442 TGGGTAGAGAAGCTGGTAACTTTTGACAGCAAGTAAACAGTTTAAAGCCATTGCTCT 1501
QY 1381 GGTCAAACTATTTCGATGATCAAAAGAAAGGAGCAACAGATTGAAACCAACCA 1440
Db 1502 GGTCAAACTATTTCGATGATCAAAAGAAAGGAGCAACAGATTGAAACCAACCA 1561
QY 1441 GTCATCATGACCAAAATGAGAACATTACAGTGGTCAAGATAGGCTGCGAAGAAAGACCA 1500
Db 1562 GTCATCATGACCAAAATGAGAACATTACAGTGGTCAAGATAGGCTGCGAAGAAAGACCA 1621
QY 1501 GAACACACTCAACCAATCAGAGACAGAACTTAACTTCACTATCACTAACACATACATTGCT 1560
Db 1622 GAACACACTCAACCAATCAGAGACAGAACTTAACTTCACTATCACTAACACATACATTGCT 1681
QY 1561 GGTGACTTTGGTTGTTGACAAATGAAATGCCCATGATTGCTGTTGGTTGTTAAAG 1620
Db 1682 GGTGACTTTGGTTGTTGACAAATGAAATGCCCATGATTGCTGTTGGTTGTTAAAG 1741
QY 1621 AATGGTTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTGATTGG 1680
Db 1742 AATGGTTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTGATTGG 1801
QY 1681 TCAGAAAATCTGGGGAGGCAAGGTGCGAACTCTTATTAATTTACTAGTTGCGGACGC 1740
Db 1802 ACGGAAAATCTGGGGAGGCAAGGTGCGAACTCTTATTAATTTACTAGTTGCGGACGC 1861
QY 1741 TCACCAATTCACGACACCGAAAAGTAGCCCTCTCAGCCAGAACTATGCACTTAACCTT 1800
Db 1862 TCACCAATTCACGACACCGAAAAGTAGCCCTCTCAGCCAGAACTATGCACTTAACCTT 1921
QY 1801 GCATCGGATCTCGAGGACCTGGCTTTTAGAGCTTTGAGACACACCAAACTCTCTGTTGCG 1860
Db 1922 GCATCGGATCTCGAGGACCTGGCTTTTAGAGCTTTGAGACACACCAAACTCTCTGTTGCG 1981
QY 1861 GGCATCTGAGAAACCCAGAACTCTGGGAGAGCTGGTTTCCAAAGCTGCGCAAGTGTCAA 1920
Db 1982 GGCATCTGAGAAACCCAGAACTCTGGGAGAGCTGGTTTCCAAAGCTGCGCAAGTGTCAA 2041
QY 1921 CTGAGCCCAATTTGGTTCAGAGATCGAGAGGATTTGAGAGCGTGTCTCGTGGGAAACCG 1980
Db 2042 CTGAGCCCAATTTGGTTCAGAGATCGAGAGGATTTGAGAGCGTGTCTCGTGGGAAACCG 2101
QY 1981 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019
```

```
Db 2102 TTGAAGAGAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2140
RESULT 14
MVU34256
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .4761
/organism="Mice minute virus"
/mol_type="genomic DNA"
/strain="Cutter"
/specific_host="Mesocricetus auratus"
/db_xref="taxon:10794"
Query Match 91.9%; Score 1855.8; DB 14; Length 4761;
Best Local Similarity 94.9%; Pred. No. 0;
Matches 1917; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
QY 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGTTAAAGGAAAAA 60
Db 122 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGTTAAAGGAAAAA 181
QY 61 AGTAACCAAGAAATGTTCTCATTTGTTTAAAAATGAAAAATGTTCAACTGAATGAAAA 120
Db 182 AGTAACCAAGTATGTTCTCATTTGTTTAAAAATGAAAGTGTTCATTTGAATGAAAA 241
QY 121 GATATCGGATGGAATGTTTACAAAAAGAGCTGCGAGGAGGAGCTGAAATCTTTACAA 180
Db 242 GATATCGGATGGAATGTTTACAAAAAGAGCTGCGAGGAGGAGCTGAAATCTTTACAA 301
QY 181 CGAGGAGCGGAACTTACTTGGGACCAAGCGAGGAGCATGGAATGGGAAACCAAGTGGAT 240
Db 302 CGAGGAGCGGAACTTACTTGGGACCAAGCGAGGAGCATGGAATGGGAAATCTTCAAGTGGAT 361
QY 241 GAAATGACCAAAAAAGCAAGTATTCTATTTTGTGTTTCTTTGTTTAAAAAATGTTTATTGAA 300
Db 362 GAACTAACCAAGCAAGTATTCTATTTGTTTGAATTTAAAGTGTTTATTGAA 421
QY 301 GTGCTTAAACAAAAAGAAATATATTTCTGTTGATGTTTAAATGTTTGTGCAACATGAATGG 360
Db 422 GTGCTAAGTACAAAAAGAAACATAGCTCTCTAGTATGTTAAATGTTGATGTCAGCATGAATGG 481
QY 361 GGAAAGACCAAGCTGCGCACTGCTCCATGTTAAATTTGGAGGAAAGCATTTAGTCAAGCT 420
Db 482 GGAAAGACCAAGCTGCGCACTGCGCAATGTTAAATTTGGAGGCAAGCACTTTAGCCAAGCT 541
QY 421 CAAGGAAATGTTGGAGAAAGGCAACTTAAATGTTTACTGAGCAGATGTTGTTGTAACAGCC 480
Db 542 CAAGGAAATGTTGGAGAAAGGCACTTAAATGTTTACTGAGCAGATGTTGTTGTAACAGCC 601
QY 481 TGTAAATGTGCAACTTAAACACAGCTGAAAGAAATTTAAACTAAGAAATAGCAGAAAGCAAT 540
```



```
Db 602 TGCAGTGTGAGCTATCACAGCGGAAAGAAATTAAGCTGAGAGAAATAGCGGAAGCCAA 661
Qy 541 GAGTGGGTACTTACTTACTTATAAGCATPAAGCAAAACCAAAAAAGAGCTATACCAAGTGT 600
Db 662 GAATGGGTCACTTTGCTTACTTATAGCAFAAGCAAAACCAAAAAAGAGCTATACCAAGTGT 721
Qy 601 GTTCTTTTGGAAACATGATGCTTACTTATTTTAACTTAAAGAAAAATAAGCACTAGT 660
Db 722 GTTGTCTTTGGAAATATGATGCTTACTTATTTTAAACCAAGAAAAATAATGCACTAGT 781
Qy 661 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTACTCTGGCTGGAAAACTTAACCTTTTAA 720
Db 782 CCACCAAGGACGAGGCTATTTTCTTAGCAGTACTCTGGCTGGAAAACTTAACCTTTTAA 841
Qy 721 AAAGAAGGAGGAGGCCATCTAGTGAGCAAACTATACATGATGACATGCGGCGAGAAACG 780
Db 842 AAAGAAGGCGAACGCCATCTAGTGAGCAAACTATACATGATGACATGCGGCGAGAAACG 901
Qy 781 GTTGAACACCACTAACCACTGCGGAGGAACTAAGCGGCGAGAACTTAACCTTAACCACTA 961
Db 902 GTTGAACACCACTAACCACTGCGGAGGAACTAAGCGGCGAGAACTTAACCTTAACCACTA 961
Qy 841 GAAGTTGCTATTAAACTACACTTAAAGAGCTGCTGCATAAAGAGTAACTCACCAGAG 900
Db 962 GAGTTTCTATTAAACCACTTAAAGAGCTGCTGCATAAAGAGTAACTCACCAGAA 1021
Qy 901 GACTGGATGATGATGACGAGCAGATTAATGAAATGATGCTCAACACAGGTGGAGAA 960
Db 1022 GACTGGATGATGATGACGAGCAGATTAATGAAATGATGCTCAGCCAGGTGGAGAA 1081
Qy 961 AACCTGCTGAAAAATACGCTAGAGATTGTACACTACTAGCCAGAACCAAAACAGCA 1020
Db 1082 AACCTGCTTAAAAATACGCTAGAGATCTGTACGCTAACTTAGCTAGAACCAAAACAGCC 1141
Qy 1021 TTTGACTTAATTTTAGAAAGCTGAACCCAGCAAACTAACCACTTTTCACTGCTGAC 1080
Db 1142 TTTGACTTAATTTTAGAAAGCTGAACCCAGCAAACTAACCACTTTTCACTGCTGAC 1201
Qy 1081 ACAAGAACCTGCAAGATTTTGTCTTTTATGCTGGAATCTATCTTAAAGTTTCCCATGCT 1140
Db 1202 ACAAGAACCTGTAAGATTTTGTCTTTTATGCTGGAATCTATCTTAAAGTTTCCCATGCT 1261
Qy 1141 ATTTGCTGTGTTTAAACAGACAAGAGGCAAAAGAAATACTGTTTATTTATGAGACCA 1200
Db 1262 ATTTGCTGTGTTTAAACAGACAAGAGGCAAAAGAAATACTGTTTATTTATGAGACCA 1321
Qy 1201 GCCAGACAGGCAATCTATTATTGACAGCCATAGCAAGCAGTTCGCAATGTTGCT 1260
Db 1322 GCCAGTACAGGCAATCCATCATTTGACAGCCATAGCAAGCAGTTCGCAATGTTGCT 1381
Qy 1261 TGCTATAATGCAAGCAATGTAACTTTTCCATTTAATGACTGTACCAACCAAGAACTTGATT 1320
Db 1382 TGCTATAATGCAAGCAATGTAACTTTTCCATTTAATGACTGTACCAACCAAGAACTTGATT 1441
Qy 1321 TGGGTAGAGAAAGCTGTAATCTTTGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT 1380
Db 1442 TGGGTAGAGAAAGCTGTAATCTTTGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT 1501
Qy 1381 GGTCAAACTATTGCAATTTGATCAAAAAGGAAAGGCAAGCAAGATTTGAAACCAACCA 1440
Db 1502 GGTCAAACTATTGCAATTTGATCAAAAAGGAAAGGCAAGCAAGATTTGAAACCAACCA 1561
Qy 1441 GTCATCATGACCAACCAATGAGAACATTACAGTGGTCAAGATAGGCTGCGAAGAAAGACCA 1500
Db 1562 GTCATCATGACCAACCAATGAGAACATTACAGTGGTCAAGATAGGCTGCGAAGAAAGACCA 1621
Qy 1501 GAACACACTCAACCAATCAGAGACAGATGCTTAAACATTCATCTAACACATACCTTTGCCT 1560
Db 1622 GAACACACTCAACCAATCAGAGACAGATGCTTAAACATTCATCTAACACATACCTTTGCCT 1681
Qy 1561 GGTGACTTTGGTTGGTGAACAAAATGAATGGCCCATGATTTGTGCTTGGTTGGTAAAG 1620
```

```
Db 1682 GGTGACTTTGGTTGGTTGACAAAAATGAATGCCCATGATTTGTGCTTGGTTGGTAAAG 1741
Qy 1621 AATGGTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCTGTATTGG 1680
Db 1742 AATGGTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCTGTATTGG 1801
Qy 1681 TCAGAAAACTGGGCGAGCCAAAGGTGCCAACTCTCTATAAATTTACTAGCTTCGGCAACG 1740
Db 1802 TCAGAAAACTGGGCGAGCCAAAGGTGCCAACTCTCTATAAATTTACTAGCTTCGGCAACG 1861
Qy 1741 TCACCAATTCACGACACCGAAAGTACGCCCTCTCAGCGAGAACTATGCACTAACTCCACTT 1800
Db 1862 TCACCAATTCAGACACCGAAAGTACGCCCTCTCAGCGAGAACTATGCACTAACTCCACTT 1921
Qy 1801 GCATCGGATCTCGAGACCTGGCTTTAGAGCCTTTAGAGCCTTTGAGACACACAAATATCTCTGTTGCG 1860
Db 1922 GCATCGGATCTCGAGACCTGGCTTTAGAGCCTTTAGAGCCTTTGAGACACACAAATATCTCTGTTGCG 1981
Qy 1861 GGCACCTGCAGAACCCAGAACACTGGGGAGCTGTTCCAAAGCCTGCCAAGATGGTCAA 1920
Db 1982 GGCACCTGCAGAACCCAGAACACTGGGGAGCTGTTCCAAAGCCTGCCAAGATGGTCAA 2041
Qy 1921 CTGAGCCCCAACTTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTCGGTGGGAACCG 1980
Db 2042 CTGAGCCCCAACTTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTCGGTGGGAACCG 2101
Qy 1981 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTTGGACTAA 2019
Db 2102 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTTGGACTAA 2140

RESULT 15
LOCUS Q0786765 5121 bp DNA linear PAT 24-MAR-2004
DEFINITION Sequence 1 from Patent WO2004018689.
ACCESSION Q0786765
VERSION Q0786765.1 GI:45721778
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Iggo, R. and Mallerba, M.
TITLE Anti-neoplastic viruses
JOURNAL Patent: WO 2004018689-A 1 04-MAR-2004;
BTG INTERNATIONAL LIMITED (GB)
FEATURES
source
1..5121
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="Parvovirus H1\WVM (phH1) - Parvovirus H1 with
promoter P4 and left hairpin from WVM"

ORIGIN

Query Match 87.7%; Score 1771; DB 6; Length 5121;
Best Local Similarity 92.3%; Pred. No. 0;
Matches 1864; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

Qy 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 60
Db 282 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 341
Qy 61 AGTAACCAAGAGTGTCTCTATTGTTTAAATAATGAAATGTTCAACTGAATGAAAA 120
Db 342 AGTAACCAAGAGTGTCTCTATTGTTTAAATAATGAAATGTTCAACTGAATGAAAA 401
Qy 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCGAGGAGCAGCTGAAATCTTTACAA 180
Db 402 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCGAGGAGCAGCTGAAATCTTTACAA 461
Qy 181 CGAGGAGCGGAAACTACTTTCGGACCAAAAGCGAGGACATGGAATGGGAAACCACTGGAT 240
```

Db 462 CGAGGAGCGGAAACTACTTGGGACCAAAAGCGAGGACATGGAATGGGAAACCAACAGTGGAT 521
Qy 241 GAAATGACCAAAAAGCAAGTATTCATTTTGGTAAATGTTTGGTAAATGTTTATTTGAA 300
Db 522 GAAATGACCAAAAAGCAAGTATTCATTTTGGTAAATGTTTGGTAAATGTTTATTTGAA 581
Qy 301 GTGCTTAACCAAAAGAAATATATTTCTGCTGATGTTAAATTTGGTTGTGCAACATGAAATGG 360
Db 582 GTGCTTAACCAAAAGAAATATATTTCTGCTGATGTTAAATTTGGTTGTGCAACATGAAATGG 641
Qy 361 GGAAGAAGCAAGCTGGCACTGTCATGTAATTTGGAGGAAAGGACTTTAGTCAAGCT 420
Db 642 GGAAGAAGCAAGCTGGCACTGTCATGTAATTTGGAGGAAAGGACTTTAGTCAAGCT 701
Qy 421 CAAGGGAATTTGGTGAAGAGCAACTAAATGTTTACTTGGAGCAGATGTTGGTAAACAGCC 480
Db 702 CAAGGGAATTTGGTGAAGAGCAACTAAATGTTTACTTGGAGCAGATGTTGGTAAACAGCC 761
Qy 481 TGTAATGTCAACTAAACACAGCTGAAAGAAATTAACCTAAGAGAAATPAGCAGAGACAAT 540
Db 762 TGTAATGTCAACTAAACACAGCTGAAAGAAATTAACCTAAGAGAAATPAGCAGAGACAAT 821
Qy 541 GAGTGGGTTACTCTACTTACTTATAGCATAAGCAACCAAAAGACTATACCAAGTGT 600
Db 822 GAGTGGGTTACTCTACTTACTTATAGCATAAGCAACCAAAAGACTATACCAAGTGT 881
Qy 601 GTTCTTTTGGAAACATGATTGCTTACTATTTTAACTAAAGAAATAAGCACTAGT 660
Db 882 GTTCTTTTGGAAACATGATTGCTTACTATTTTAACTAAAGAAATAAGCACTAGT 941
Qy 661 CCACCAAGAGACGGAGGCTATTTCTTAGCAGTACTCTGGCTGGAAACATACTTTTAA 720
Db 942 CCACCAAGAGACGGAGGCTATTTCTTAGCAGTACTCTGGCTGGAAACATACTTTTAA 1001
Qy 721 AAGAGGCGAGCGCCATCTAGTGAGCAAACTATACCTGATGACATGGGCCAGAAACG 780
Db 1002 AAGAGGCGAGCGCCATCTAGTGAGCAAACTATACCTGATGACATGGGCCAGAAACG 1061
Qy 781 GTTGAACACACAGTAACCACTGGCGAGAACTAAGCGCGGAGAAATTCAAACTAAAAAA 840
Db 1062 GTCGAGNACACAGTGACCACTGCACAGGAGCTAAGCGCGGAGAAATTCAAACTAGAGAG 1121
Qy 841 GAAGTTGCTATTTAAACCTACATTAAGAGCTGGTGCAATAAAGAGTAACCTCACAGAG 900
Db 1122 GAGGCTCTCGATTAAACCACTCAAGAGTTGGTACATAAAGAGTAACCTCACAGAA 1181
Qy 901 GACTGGATGATGATGAGCAGACAGTTACATTTGAATGATGCTCAACCGAGTGGAGAA 960
Db 1182 GACTGGATGATGATGAGCAGACAGTTACATTTGAATGATGCTCAACCGAGTGGAGAA 1241
Qy 961 AACCTGTGAAAAATACGCTAGAGATTGTACATACTAGCCAGAAACCAAAACAGCA 1020
Db 1242 AACCTGTGAAAAATACGCTAGAGATTGTACATACTAGCCAGAAACCAAAACAGCC 1301
Qy 1021 TTTGACTTAATTTTGAAGAAAGCTGAAAACAGCAAACTAACCACTTTTCACTGGCTGAC 1080
Db 1302 TTTGACTTGAATTTGGAAGAAAGCTGAAAACAGCAAACTAACCACTTTTCACTGGCTGAC 1361
Qy 1081 ACAAGAACCTGCAAGATTTTGGCTTTTTCATGGCTGGAACTATGTTAAAGTTGCCATGCT 1140
Db 1362 ACCAGAACCTGTAAGATCTTTGCTGAGCATGGCTGGAACTATGTTAAAGTTGCCATGCT 1421
Qy 1141 ATTGTGCTGTTTTTAAACAGACAGGAGGCAAAAGAAATACTGTTTTTATTTTCAATGACCA 1200
Db 1422 ATCTGTTGTGCTGAATAGACAGGAGGCAAAAGGAACACTGTGCTCTTTTCAAGCA 1481
Qy 1201 GCCAGACAGGCAAAATCTATTATTGCAAGCCATAGCAACAGCTTTGGCAATGTTGTT 1260
Db 1482 GCCAGACAGGCAAAATCTATTATTGCAAGCCATAGCAACAGCTTTGGCAATGTTGTT 1541
Qy 1261 TGCTATAATGACGCAATGTAACCTTTTCCATTTTAAATGACTGTACCAACAGAACTTTGATT 1320
Db 1542 TGTTACATGCTGCCAATGTAACCTTTCCATTTTAAATGACTGTACCAACAAAAACTTTGATT 1601

Qy 1321 TGGGTAGAAGAGCTGCTAACTTTTGACAGCAAGTAAACCCAGTTTAAAGCCATTGCTCT 1380
Db 1602 TGGGTGGAAGAGCTGCTAACTTTTGCCAGCAAGTAAACCAATTCAGAGCTATTGTTCT 1661
Qy 1381 GGTCAAACTATTTCGCATTGATCAAAAGGAAAGGAGCAAAAGGAGCAATTCAGCAACCA 1440
Db 1662 GGCCAAACCATAGCATTGATCAAAAGGAAAGGAGCAAAAGGAGCAATTCAGCAACCA 1721
Qy 1441 GTCATCATGACCAACATGAGAACATTACAGTGGTTCAGATAGGCTTCGGAAGAAAGCA 1500
Db 1722 GTTATTATGACCAACCAACGAGAACATTACCGTGGTTAGAAATAGGCTGTGAGGAAGACCA 1781
Qy 1501 GAAACACATCAACCAATCAGAGACAGAAATGCTTAAACATTCATCTAACACATACCTTTGCCCT 1560
Db 1782 GAAACACATCAACCAATCAGAGACAGAAATGCTTAAACATTCATCTAACACATACCTTTGCCCT 1841
Qy 1561 GGTGACTTTGGTTGGTTGCACAAAATGAATGGCCCATGATTTGTCTTTGGTTGGTAAAG 1620
Db 1842 GGTGACTTTGGTTGGTTGCATGAATGGCCCTCTGATCTGTGCTTTGGTTGGTGAAG 1901
Qy 1621 AATGGTTACCAATCTACCATGGCAAGTACTGCTTAAATGGGCAAAAGTTCCTGATTGG 1680
Db 1902 AATGGTTACCAATCTACCATGGCTTTGTTACTGTGCTTAAATGGGCAAAAGTTCCTGATTGG 1961
Qy 1681 TCAGAAAACTGGCGGAGGCAAAAGGTGCCAACTCCTATAAAATTTACTAGGTTGGGCACGC 1740
Db 1962 TCAGAGGACTGGCGGAGGCGGAAGCTAGACACTCCTATAAATTCGCTAGGTTCAATGGCG 2021
Qy 1741 TCACCAATTCACACACCGAAAGTAGCCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800
Db 2022 TCACCAATCTCAGCTCCGAGAAAGTACGCTCTCAGCCAAATCTAGCTTTACTCCACTT 2081
Qy 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTGAGCAGACCAACCAATACTCCTGTTGCG 1860
Db 2082 GCATCGGACCTTTGCGGACCTAGCTCTAGAGCCTTGAGCAGACCAACCAATACTCCTGTTGCG 2141
Qy 1861 GGCACCTGCAGAAAAACCCAGAAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGCTCAA 1920
Db 2142 GGCACCTGCAGAAAGCCAAACACTGGGAGGCTGGTTCCACAGCCTGCCAAGTGTCTCAA 2201
Qy 1921 CTGAGCCCAACTTTGGTTCAGAGATTCGAGGAGGATTTTGAGAGCGTGTCTTGGTGGGAAACCG 1980
Db 2202 CGGAGCCCAACCTGGTTCGAGATCGAGGCGGATTTGAGAGCTTTGCTTCAGTCAAGAAACAG 2261
Qy 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTTGGACTAA 2019
Db 2262 TTGGAGAGGACTTCAACGAGGAGCTGACCTTTGGACTAA 2300

Search completed: January 22, 2005, 21:19:15
Job time : 8548.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2005, 23:46:57 ; Search time 950.5 Seconds
(without alignments)
11150.532 Million cell updates/sec

Title: us-10-069-056-4
Perfect score: 2019
Sequence: 1 atggcgtggaatgcttactc.....agccgtgaactggactaa 2019

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2019	100.0	2019	5	AAD02799
2	2017.4	99.9	2019	5	AAD02797
3	2017.4	99.9	5149	12	ADG39767
4	2015.8	99.8	2019	5	AAD02801
5	2015.8	99.8	2019	5	AAD02805
6	2015.8	99.8	2019	5	AAD02803
7	1040.4	51.5	5049	2	AAT15311
8	1040.4	51.5	5049	2	AAT15312
9	1040.4	51.5	5049	2	AAT88324
10	1038.8	51.5	5049	2	AAT88321
11	1038.8	51.5	5049	2	AAT88320
12	902.6	44.7	3524	1	AAN40252
13	307.4	15.2	374	6	ABQ95626
14	285.8	14.2	421	6	ABQ94779
15	284.8	14.1	423	6	ABQ94724
16	284.8	14.1	464	6	ABQ94719
17	284.8	14.1	473	6	ABQ94716
18	284.8	14.1	486	6	ABQ94707
19	284.8	14.1	491	6	ABQ94777
20	284.2	14.1	420	6	ABQ94787
21	283.8	14.1	343	6	ABQ94737

22	283.8	14.1	343	6	ABQ94739	Abq94739 Tumour su
23	283.8	14.1	403	6	ABQ94782	Abq94782 Tumour su
24	283.2	14.0	497	6	ABQ94718	Abq94718 Tumour su
25	280.4	13.9	340	6	ABQ94732	Abq94732 Tumour su
26	278.2	13.8	342	6	ABQ94765	Abq94765 Tumour su
27	273.8	13.6	420	6	ABQ94781	Abq94781 Tumour su
28	269.6	13.4	324	6	ABQ94740	Abq94740 Tumour su
29	263.4	13.0	339	6	ABQ95625	Abq95625 Tumour su
30	261.6	13.0	424	6	ABQ94778	Abq94778 Tumour su
31	163.8	8.1	468	6	ABQ94873	Abq94873 Tumour su
32	162.2	8.0	451	6	ABQ94868	Abq94868 Tumour su
33	162.2	8.0	457	6	ABQ94867	Abq94867 Tumour su
34	162.2	8.0	465	6	ABQ94874	Abq94874 Tumour su
35	162.2	8.0	469	6	ABQ94869	Abq94869 Tumour su
36	162.2	8.0	472	6	ABQ94880	Abq94880 Tumour su
37	162.2	8.0	515	6	ABQ94793	Abq94793 Tumour su
38	162.2	8.0	516	6	ABQ94807	Abq94807 Tumour su
39	162.2	8.0	516	6	ABQ94794	Abq94794 Tumour su
40	162.2	8.0	516	6	ABQ94800	Abq94800 Tumour su
41	162.2	8.0	517	6	ABQ94811	Abq94811 Tumour su
42	162.2	8.0	530	6	ABQ94791	Abq94791 Tumour su
43	162.2	8.0	530	6	ABQ94799	Abq94799 Tumour su
44	161	8.0	516	6	ABQ94796	Abq94796 Tumour su
45	160.6	8.0	530	6	ABQ94805	Abq94805 Tumour su

ALIGNMENTS

RESULT 1

ID AAD02799 standard; DNA; 2019 BP.
XX
AC AAD02799;
XX
DT 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
DE Parvovirus non-structure protein 1 (NS1) variant (S283A) DNA.
XX
KW NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
tumoural disease; gene therapy; mutant; mutein; variant; ds.
XX
OS Parvovirus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS i..2019
FT /*tag= a
FT /*product= "NS1 variant (S283A) protein"
FT /*replace(847, A)
FT /*tag= b
XX
FT mutation
XX
FT EPI077260-AL.
XX
PD 21-FEB-2001.
XX
PF 13-AUG-1999; 99EP-00115161.
XX
PR 13-AUG-1999; 99EP-00115161.
XX
(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
PI Nueesch J, Rommelaere J;
XX
DR WPI; 2001-212717/22.
DR P-PSDB; AAY72704.
XX
PT Novel parvovirus non-structure protein variant, useful for treating
tumoral diseases, has a shifted equilibrium between DNA replication and
transcription activities, and cytotoxic activity.
XX
PS Claim 7; Page 11-14; 41pp; English.

XX	The present sequence is a DNA encoding parvovirus non-structure 1 protein (NS1) variant (S283A). The invention relates to the variants of the parvovirus non-structure protein (NS1) having a shifted equilibrium between the DNA replication and transcription activities, and the cytotoxicity activity. These variants are useful as toxins for treating tumoural diseases. The variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-2003 to correct OS field.)									
XX	Sequence 2019 BP; 698 A; 389 C; 453 G; 479 T; 0 U; 0 Other;									
QY	Query Match	100.0%;	Score 2019;	DB 5;	Length 2019;					
CC	Best Local Similarity	100.0%;	Pred. No. 0;							
CC	Matches 2019; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;					
QY	1	ATGGCTGGAAATGCTTACTCTGATGAGTTTGGGAGCAACCAACTGGTTAAAGGAAAAA	60							
Dd	1	ATGGCTGGAAATGCTTACTCTGATGAGTTTGGGAGCAACCAACTGGTTAAAGGAAAAA	60							
QY	61	AGTAACGAGGAAGTGTCTCATTTGTTTTTAAAAATGAAAAATGTTCAACTGAATGAAAA	120							
Dd	61	AGTAACGAGGAAGTGTCTCATTTGTTTTTAAAAATGAAAAATGTTCAACTGAATGAAAA	120							
QY	121	GATATCGGATGGAATAGTTACAAAAAGAGCTCGAGGAGCAAGCTGTAATCTTTACAA	180							
Dd	121	GATATCGGATGGAATAGTTACAAAAAGAGCTCGAGGAGCAAGCTGTAATCTTTACAA	180							
QY	181	CGAGGAGCGGAATCTACTTTGGACCAAGCGAGGACATGGAATGGAAACCAACAGTGGAT	240							
Dd	181	CGAGGAGCGGAATCTACTTTGGACCAAGCGAGGACATGGAATGGAAACCAACAGTGGAT	240							
QY	241	GAAATGACCAAAAAAGCAAGTATTCAATTTTGTATTTTGGTTAAAAAATGTTTATTGAA	300							
Dd	241	GAAATGACCAAAAAAGCAAGTATTCAATTTTGTATTTTGGTTAAAAAATGTTTATTGAA	300							
QY	301	GTGCTTAAACAAGAATATATTCTCTGTGATGTTAAATGTTTGTGCAACATGAATGG	360							
Dd	301	GTGCTTAAACAAGAATATATTCTCTGTGATGTTAAATGTTTGTGCAACATGAATGG	360							
QY	361	GGAAAGACCAAGCTGGCACTGCCATGATTAATTTGGGAAAGACATTTAGTCAAGCT	420							
Dd	361	GGAAAGACCAAGCTGGCACTGCCATGATTAATTTGGGAAAGACATTTAGTCAAGCT	420							
QY	421	CAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGGTAACAGCC	480							
Dd	421	CAAGGGAATGTTGGAGAGGCACTAAATGTTTACTGGAGCAGATGGTTGGTAACAGCC	480							
QY	481	TGTAATGTGCAACTAACAACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAACAT	540							
Dd	481	TGTAATGTGCAACTAACAACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAACAT	540							
QY	541	GAGTGGGTTACTCTACTTATTAAGCATAAGCAAAACCAAAAAGACTATACCAAGTGT	600							
Dd	541	GAGTGGGTTACTCTACTTATTAAGCATAAGCAAAACCAAAAAGACTATACCAAGTGT	600							
QY	601	GTTCTTTTGGAAACATGATTGCTTACTATTTTTTAACTAAAAAGAAAAATTAAGCACTAGT	660							
Dd	601	GTTCTTTTGGAAACATGATTGCTTACTATTTTTTAACTAAAAAGAAAAATTAAGCACTAGT	660							
QY	661	CCAACAGAGACGGAGGCTATTTTCTTAGCAGTGAATCTGGCTGGAAAACTAACTTTTAA	720							
Dd	661	CCAACAGAGACGGAGGCTATTTTCTTAGCAGTGAATCTGGCTGGAAAACTAACTTTTAA	720							
QY	721	AAAGAGGGCGAGCGCCATCTAGTGAGCAAACTATACATGATCAGATGGGCCAGAAACG	780							
Dd	721	AAAGAGGGCGAGCGCCATCTAGTGAGCAAACTATACATGATCAGATGGGCCAGAAACG	780							
QY	781	GTTGAAACCAACAGTAACCTGCGCAGGAAACTAAGCGCGGCAGAAATTCAAACTAAAAA	840							
Dd	781	GTTGAAACCAACAGTAACCTGCGCAGGAAACTAAGCGCGGCAGAAATTCAAACTAAAAA	840							
QY	841	GAAATGCTATTAAAACTACATTAAGAGCTGGTGATTAATAAGAGTAGTAACCTCACAGAG	900							

Qy	181	CGAGGAGCGGAAACTTACTTTGGGACCAAGCGAGGACATGGAAATGGGAAACCAACAGTGGAT	240
Db	181	CGAGGAGCGGAAACTTACTTTGGGACCAAGCGAGGACATGGAAATGGGAAACCAACAGTGGAT	240
Qy	241	GAATGACCAAAAAAGCAAGTATTCTTTTGGTTAAATAATGTTTATTGAA	300
Db	241	GAATGACCAAAAAAGCAAGTATTCTTTTGGTTAAATAATGTTTATTGAA	300
Qy	301	GTGCTTTAAACAACAAGAAATATATTTCTCTGTTGATGTTAAATGGTTTGTGCAACATGAATGG	360
Db	301	GTGCTTTAAACAACAAGAAATATATTTCTCTGTTGATGTTAAATGGTTTGTGCAACATGAATGG	360
Qy	361	GGAAAAGACCAAGCTGGCACTGCCATGTACTTAATTTGGAGAAAGGACTTTAGTCAAGCT	420
Db	361	GGAAAAGACCAAGCTGGCACTGCCATGTACTTAATTTGGAGAAAGGACTTTAGTCAAGCT	420
Qy	421	CAAGGGAATGTTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGTTGGTAAACAGCC	480
Db	421	CAAGGGAATGTTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGTTGGTAAACAGCC	480
Qy	481	TGTAATGTGCAACTAACACCACTGAAAGAAATTAACCTAAGAGAAATAGCAGAACAAT	540
Db	481	TGTAATGTGCAACTAACACCACTGAAAGAAATTAACCTAAGAGAAATAGCAGAACAAT	540
Qy	541	GAGTGGTTACTCTTACTTATAGCATAAGCAAAACCAAAAGACTATACCAAGTGT	600
Db	541	GAGTGGTTACTCTTACTTATAGCATAAGCAAAACCAAAAGACTATACCAAGTGT	600
Qy	601	GTTCTTTTGGAAACATGATTGCTTACTATTTTAACTAAAGAAATTAAGCACTAGT	660
Db	601	GTTCTTTTGGAAACATGATTGCTTACTATTTTAACTAAAGAAATTAAGCACTAGT	660
Qy	661	CCACCAAGAGACGAGGCTATTTTCTAGCAGTACTCTGCTGGGAAAACTAACTTTTAA	720
Db	661	CCACCAAGAGACGAGGCTATTTTCTAGCAGTACTCTGCTGGGAAAACTAACTTTTAA	720
Qy	721	AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG	780
Db	721	AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG	780
Qy	781	GTTTGAACACACAGTAAACCACTGCGAGGAACTTAAGCGCGGAGAAATTAACCTAAAAA	840
Db	781	GTTTGAACACACAGTAAACCACTGCGAGGAACTTAAGCGCGGAGAAATTAACCTAAAAA	840
Qy	841	GAAAGTTGCTATTAAAACTACACTTAAAGAGCTGGTGATTAAGAGAGTAACTCACCAGAG	900
Db	841	GAAAGTTGCTATTAAAACTACACTTAAAGAGCTGGTGATTAAGAGAGTAACTCACCAGAG	900
Qy	901	GACTGATGATGATGAGCGACAGTTACATTGAAATGATGCTCAACGAGTGGAGAA	960
Db	901	GACTGATGATGATGAGCGACAGTTACATTGAAATGATGCTCAACGAGTGGAGAA	960
Qy	961	AACCTGCTGAAAAATAGCTTAGAGATTTGTACACTAATCTAGCCAGAACCAAAACAGCA	1020
Db	961	AACCTGCTGAAAAATAGCTTAGAGATTTGTACACTAATCTAGCCAGAACCAAAACAGCA	1020
Qy	1021	TTTGACTTAAATTTAGAAAAAGCTGAAACCCAGCAAACTAACCACTTTTCACTGCCTGAC	1080
Db	1021	TTTGACTTAAATTTAGAAAAAGCTGAAACCCAGCAAACTAACCACTTTTCACTGCCTGAC	1080
Qy	1081	ACAAGAACCTGCAAGATTTTGTCTTTTCACTGGTGAACATATGTTAAAGTTTGGCATGCT	1140
Db	1081	ACAAGAACCTGCAAGATTTTGTCTTTTCACTGGTGAACATATGTTAAAGTTTGGCATGCT	1140
Qy	1141	ATTGCTGTGTTTAAACAGAGGAGGCAAAAGAAATCTGTTTATTTATGACGACCA	1200
Db	1141	ATTGCTGTGTTTAAACAGAGGAGGCAAAAGAAATCTGTTTATTTATGACGACCA	1200
Qy	1201	GCCAGACAGGCAAAATCTATTATTGCAACAGCCATAGCAAGCAGTGGCAATGTTGGT	1260
Db	1201	GCCAGACAGGCAAAATCTATTATTGCAACAGCCATAGCAAGCAGTGGCAATGTTGGT	1260

Qy	1981	TTGAAGAAAGACTTTCAGGAGCCGCTGAACTTGACATAA	2019
Db	1981	TTGAAGAAAGACTTTCAGGAGCCGCTGAACTTTGGACTAA	2019
RESULT 2			
AAD02797 standard; DNA; 2019 BP.			
ID	AC	AAD02797;	
XX	AC		
XX	06-AUG-2003	(revised)	
DT	31-MAY-2001	(first entry)	
XX	Parvovirus non-structure protein 1 (NS1) wild-type DNA.		
DE	NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;		
KW	tumoural disease; gene therapy; ds.		
XX	Parvovirus.		
OS			
FH	Key	Location/Qualifiers	
FT	CDS	1..2019	
FT	/*tag= a		
FT	/product= "Parvovirus NS1 protein"		
XX	EP1077260-A1.		
PN	XX		
XX	21-FEB-2001.		
PD	XX		
XX	13-AUG-1999;	99EP-00115161.	
PF	XX		
PR	13-AUG-1999;	99EP-00115161.	
XX	(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.		
PA			
XX	Nuesch J, Rommelaere J;		
PI	WPI; 2001-212717/22.		
XX	P-PSDB; AAY72702.		
DR	Novel parvovirus non-structure protein variant, useful for treating		
XX	tumoral diseases, has a shifted equilibrium between DNA replication and		
PT	transcription activities, and cytotoxic activity.		
PT	Disclosure; Fig 1; 4lpp; English.		
XX	The present sequence is a wild type DNA encoding parvovirus non-		
CC	structure protein 1 (NS1). The present invention relates to the variants		
CC	of the parvovirus non-structure protein (NS1) having a shifted		
CC	equilibrium between the DNA replication and transcription activities, and		
CC	the cytotoxicity activity. These variants are useful as toxins for		
CC	treating tumoural diseases. The variant DNAs are useful as vectors for		
CC	gene therapy. (Updated on 06-AUG-2003 to correct OS field.)		
XX			
PS	Sequence 2019 BP; 698 A; 389 C; 452 G; 480 T; 0 U; 0 Other;		
SQ			
Query Match 99.9%; Score 2017.4; DB 5; Length 2019;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 2018; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	1	ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGTTAAAGGAAAA	60
Db	1	ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGTTAAAGGAAAA	60
Qy	61	AGTAACAGGAAGTGTCTCATTTTAAAAATGAAATGTTCAACTGAATGGAAAA	120
Db	61	AGTAACAGGAAGTGTCTCATTTTAAAAATGAAATGTTCAACTGAATGGAAAA	120
Qy	121	GATATCGGATGGATAGTTACAAAAAGAGCTGCAGGACGAGCTGAAATCTTTACAA	180
Db	121	GATATCGGATGGATAGTTACAAAAAGAGCTGCAGGACGAGCTGAAATCTTTACAA	180

```
QY 1261 TGCTATAATGACGCCAATGTAAACTTTCCATTTAATGACTGTACCAACCAAGAACTTGATT 1320
Db 1261 TGCTATAATGACGCCAATGTAAACTTTCCATTTAATGACTGTACCAACCAAGAACTTGATT 1320
QY 1321 TGGGTAGAGAAGCTGGTAACCTTTTGACAGCAAGTAAACCAAGTTTAAAGCCATTGCTCT 1380
Db 1321 TGGGTAGAGAAGCTGGTAACCTTTTGACAGCAAGTAAACCAAGTTTAAAGCCATTGCTCT 1380
QY 1381 GGTCAAACTATTTCGCATTGATCAAAAGGAAAGGAGGAGCAAGATTCAGATTCAGCAACCA 1440
Db 1381 GGTCAAACTATTTCGCATTGATCAAAAGGAAAGGAGGAGCAAGATTCAGATTCAGCAACCA 1440
QY 1441 GTCATCATGACCAACCAATGAGAACTTACAGTGGTCAAGATAGGCTGCGAAGAAAGACCA 1500
Db 1441 GTCATCATGACCAACCAATGAGAACTTACAGTGGTCAAGATAGGCTGCGAAGAAAGACCA 1500
QY 1501 GAACACATCAACCAATCAGAGACAGATGCTTAACATTCATCATACATACATACCTTGCTT 1560
Db 1501 GAACACATCAACCAATCAGAGACAGATGCTTAACATTCATCATACATACATACCTTGCTT 1560
QY 1561 GGTGACTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGCTTGGTTGGTAAAG 1620
Db 1561 GGTGACTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGCTTGGTTGGTAAAG 1620
QY 1621 AATGGTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCCTGATTGG 1680
Db 1621 AATGGTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCCTGATTGG 1680
QY 1681 TCAGAAAAGTGGGGAGGAGCAAGGTCGCAACTCTTAAATTTACTAGTTGGCGACGC 1740
Db 1681 TCAGAAAAGTGGGGAGGAGCAAGGTCGCAACTCTTAAATTTACTAGTTGGCGACGC 1740
QY 1741 TCACCAATTCAGACACCGAAAAGTACGCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800
Db 1741 TCACCAATTCAGACACCGAAAAGTACGCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800
QY 1801 GCATCGGATCTCAGGACCTGGCTTTAGAGCTTTGAGAGCACCAAACTCTCTGTTGG 1860
Db 1801 GCATCGGATCTCAGGACCTGGCTTTAGAGCTTTGAGAGCACCAAACTCTCTGTTGG 1860
QY 1861 GGCATCTGCAAAACCCAGACACTGGGAGAGCTGTTCCAAAGCCTGCCAAGATGTCAA 1920
Db 1861 GGCATCTGCAAAACCCAGACACTGGGAGAGCTGTTCCAAAGCCTGCCAAGATGTCAA 1920
QY 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTGGTGGGAACCG 1980
Db 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTGGTGGGAACCG 1980
QY 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2019
Db 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2019
```

RESULT 3

```
ADG39767
ID ADG39767 standard; DNA; 5149 BP.
XX
AC ADG39767;
XX
XX 11-MAR-2004 (first entry)
XX
DE Minute virus from mouse genomic DNA sequence SEQ ID NO:29.
XX
XX parvovirus; rep; cap; DNA binding domain; capsid interacting domain;
KW recombinant hybrid parvovirus particle;
KW recombinant adeno-associated virus; rAAV; AAV; vaccine; gene therapy;
KW gene; ds.
XX
OS Mice minute virus.
XX
XX WO2003104392-A2.
XX
PD 18-DEC-2003.
```

```
XX
PF 02-DEC-2002; 2002WO-US038423.
XX
PR 18-DEC-2001; 2001US-0341919P.
XX
PA (UYNC-) UNIV NORTH CAROLINA.
XX
PI Samulski RJ, Rabinowitz JE;
XX
XX WPI; 2004-062324/06.
XX
PT New polynucleotides comprising parvovirus rep coding sequences and
PT parvovirus cap coding sequences, useful in producing higher stocks of
PT hybrid parvovirus vectors for delivering therapeutic nucleic acids to a
PT subject.
XX
PS Disclosure; SEQ ID NO 29; 115pp; English.
XX
CC The present invention describes a polynucleotide (I), comprising
CC parvovirus rep coding sequences and parvovirus cap coding sequences. The
CC rep coding sequences encodes a DNA binding domain from a first
CC parvovirus, and a capsid interacting domain from a parvovirus different
CC from the first parvovirus. The cap coding sequence comprises sequences
CC from the different parvovirus. Also described: (1) a vector comprising
CC (1); (2) a cell comprising (1), or parvovirus rep coding sequences and
CC parvovirus cap coding sequences, where the rep coding sequences encode a
CC DNA binding domain from a first parvovirus and a capsid interacting
CC domain from a parvovirus different from the first parvovirus, the cap
CC coding sequences comprise sequences from the different parvovirus, and
CC the rep coding sequences are stably integrated into the genome of the
CC cell; and (3) producing a recombinant hybrid parvovirus particle or adeno
CC -associated virus (rAAV) particle. (I) can be used in vaccines, and in
CC gene therapy. The polynucleotide (I) can be used in producing higher
CC stocks of hybrid parvovirus or parvovirus vectors, which may be used in
CC the delivery of nucleic acids having biological effect to treat or
CC ameliorate the symptoms associated with any disorder related to gene
CC expression. The polynucleotide may be used to produce a parvovirus vector
CC to express an immunogenic polypeptide in a subject, e.g. for vaccination.
CC The parvovirus vector may also be used to provide an antisense nucleic
CC acid to a cell in vitro or in vivo, or in diagnostic and screening
CC methods. The present sequence is used in the exemplification of the
CC present invention.
XX
SQ Sequence 5149 BP; 1718 A; 1045 C; 1124 G; 1262 T; 0 U; 0 Other;
Query Match 99.9%; Score 2017.4; DB 12; Length 5149;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2018; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 60
Db 261 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 320
QY 61 AGTAACCAAGAACTGTTCTCATTTGTTTAAAAATGAAAAATGTTCAACTGAATGAAAA 120
Db 321 AGTAACCAAGAACTGTTCTCATTTGTTTAAAAATGAAAAATGTTCAACTGAATGAAAA 380
QY 121 GATATCGGATGGAATAGTTTCAAAAAAGAGCTGCAGAGGAGCAGAGCTGAATCTTTACAA 180
Db 381 GATATCGGATGGAATAGTTTCAAAAAAGAGCTGCAGAGGAGCAGAGCTGAATCTTTACAA 440
QY 181 CGAGGAGCGGAACACTTCTGGGACCAAGCGGACATGGAATGGGAACCAACGATGGAT 240
Db 441 CGAGGAGCGGAACACTTCTGGGACCAAGCGGACATGGAATGGGAACCAACGATGGAT 500
QY 241 GAAATGACCAAAAGCAAGTATTTCATTTTGGTTGTTTGGTTAAAAAATGTTTATTGAA 300
Db 501 GAAATGACCAAAAGCAAGTATTTCATTTTGGTTGTTTGGTTAAAAAATGTTTATTGAA 560
QY 301 GTGCTTTAACACAAAGAAATATATTTCTCGTGTGATGTTAATTTGTTTGCACATGATGG 360
Db 561 GTGCTTTAACACAAAGAAATATATTTCTCGTGTGATGTTAATTTGTTTGCACATGATGG 620
```



```
Qy 361 GGAAGACCAAGGCTGCACTGCCATCTACTAATTGGAGGAGGACTTTAGTCAAGCT 420
Db 621 GGAAGACCAAGGCTGCACTGCCATCTACTAATTGGAGGAGGACTTTAGTCAAGCT 680
Qy 421 CAAGGGAATGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTGGTAACAGCC 480
Db 681 CAAGGGAATGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTGGTAACAGCC 740
Qy 481 TGTAAATGTGAACTAAACACAGCTGAAAGAAATTAACCTAAGAGAAAATAGCAGAAACAAT 540
Db 741 TGTAAATGTGAACTAAACACAGCTGAAAGAAATTAACCTAAGAGAAAATAGCAGAAACAAT 800
Qy 541 GAGTGGGTACTCTACTTACTTATAAGCATAGCAACCAACCAAAAGACTATACCAAGTGT 600
Db 801 GAGTGGGTACTCTACTTACTTATAAGCATAGCAACCAACCAAAAGACTATACCAAGTGT 860
Qy 601 GTTCTTTTGGAAACATGATTTGCTTACTATTTTTTAACTAAAGAGAAAATAAGCACTAGT 660
Db 861 GTTCTTTTGGAAACATGATTTGCTTACTATTTTTTAACTAAAGAGAAAATAAGCACTAGT 920
Qy 661 CCACCAAGAGACGGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACCTAACTTTTA 720
Db 921 CCACCAAGAGACGGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACCTAACTTTTA 980
Qy 721 AAAGAGCGGCGGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG 780
Db 981 AAAGAGCGGCGGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG 1040
Qy 781 GTTGAACACAGTAACCACTGCGCAGGAACTAAGCGCGCAGAAATTCAAACTAAAAA 840
Db 1041 GTTGAACACAGTAACCACTGCGCAGGAACTAAGCGCGCAGAAATTCAAACTAAAAA 1100
Qy 841 GAAGTTGCTATTAAAACTACATTTAAAGAGCTGGTGCATAAAAAGAGTAACCTACCCAGAG 900
Db 1101 GAAGTTTCTATTAAAACTACATTTAAAGAGCTGGTGCATAAAAAGAGTAACCTACCCAGAG 1160
Qy 901 GACTGGATGATGTCGACGACAGTACATTGAATGATGCTCAACCAAGTGGAGAA 960
Db 1161 GACTGGATGATGTCGACGACAGTACATTGAATGATGCTCAACCAAGTGGAGAA 1220
Qy 961 AACTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAAACCAAAACAGCA 1020
Db 1221 AACTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAAACCAAAACAGCA 1280
Qy 1021 TTTGACTTAAATTTTAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCTGAC 1080
Db 1281 TTTGACTTAAATTTTAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCTGAC 1340
Qy 1081 ACAAGAACCTGCAGAAATTTTGTCTTTTCATGGCTGGAACTATGTTAAAGTTTGCATGCT 1140
Db 1341 ACAAGAACCTGCAGAAATTTTGTCTTTTCATGGCTGGAACTATGTTAAAGTTTGCATGCT 1400
Qy 1141 ATTTGCTGTGTTTTTAAACAGACAGAGGAGGCAAAAGAAATACTGTTTTATTTCATGGACCA 1200
Db 1401 ATTTGCTGTGTTTTTAAACAGACAGAGGAGGCAAAAGAAATACTGTTTTATTTCATGGACCA 1460
Qy 1201 GCCAGCACAGGCAAACTATTATTGCAACAGCCATAGCACAAGCAGTTGGCAATGTTGGT 1260
Db 1461 GCCAGCACAGGCAAACTATTATTGCAACAGCCATAGCACAAGCAGTTGGCAATGTTGGT 1520
Qy 1261 TGCTATAATGAGCCAAATCTATTATTGCAACAGCCATAGCACAAGCAGTTGGCAATGTTGGT 1320
Db 1521 TGCTATAATGAGCCAAATCTATTATTGCAACAGCCATAGCACAAGCAGTTGGCAATGTTGGT 1580
Qy 1321 TGGGTAGAGAGCTGCTTAACCTTTGGACAGCAGTAAACAGTTTAAAGCCATTGCTCT 1380
Db 1581 TGGGTAGAGAGCTGCTTAACCTTTGGACAGCAGTAAACAGTTTAAAGCCATTGCTCT 1640
Qy 1381 GGTCAAACTATTGCAATTTGATCAAAAAGAAAAGGAGGCAAAACAGATTGAAACCAACCA 1440
Db 1641 GGTCAAACTATTGCAATTTGATCAAAAAGAAAAGGAGGCAAAACAGATTGAAACCAACCA 1700
Qy 1441 GTCATCATGACCAAAATGAGAACTTACAGTGGTTCAGAAATAGGCTCGGAAGAAAGCA 1500
```

```
Db 1701 GTCATCATGACCAAAATGAGAACTTACAGTGGTTCAGAAATAGGCTCGGAAGAAAGCA 1760
Qy 1501 GAACACACTCAACCAATCAGAGACAGAAATGCTTAAACATTTCACTTAACACATACCTTGCT 1560
Db 1761 GAACACACTCAACCAATCAGAGACAGAAATGCTTAAACATTTCACTTAACACATACCTTGCT 1820
Qy 1561 GGTGACTTTGGTGGTGGTGAACAAATGAATGGCCCAATGATTTGTGCTTGGTGGTAAAG 1620
Db 1821 GGTGACTTTGGTGGTGGTGAACAAATGAATGGCCCAATGATTTGTGCTTGGTGGTAAAG 1880
Qy 1621 AATGTTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCCCTGATTGG 1680
Db 1881 AATGTTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCCCTGATTGG 1940
Qy 1681 TCAGAAAACCTGGGCGGAGCCAAAGGTGCCAACTCCTATAAATTTTACTAGGTTCGGCAGCG 1740
Db 1941 TCAGAAAACCTGGGCGGAGCCAAAGGTGCCAACTCCTATAAATTTTACTAGGTTCGGCAGCG 2000
Qy 1741 TCACCAATTCAGACACACGAAAGATAGCGCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800
Db 2001 TCACCAATTCAGACACACGAAAGATAGCGCTCTCAGCCAGAACTATGCACTAACTCCACTT 2060
Qy 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAAACTACTCCTGTTGCG 1860
Db 2061 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAAACTACTCCTGTTGCG 2120
Qy 1861 GGCACCTGCAGAAACCCAGAAACACCTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGGTCAA 1920
Db 2121 GGCACCTGCAGAAACCCAGAAACACCTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGGTCAA 2180
Qy 1921 CTGAGCCCAACTGGTTCAGAGATCGAGAGGATTTGAGAGCGTCTTCGGTTCGGGAACCG 1980
Db 2181 CTGAGCCCAACTGGTTCAGAGATCGAGAGGATTTGAGAGCGTCTTCGGTTCGGGAACCG 2240
Qy 1981 TTGAAGAAACACTTCAGCGAGCGCTGAACTTCGACTAA 2019
Db 2241 TTGAAGAAACACTTCAGCGAGCGCTGAACTTCGACTAA 2279

RESULT 4
AAD02801
ID AAD02801 standard; DNA; 2019 BP.
XX
AC AAD02801;
XX AC
XX 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
DE Parvovirus non-structure protein 1 (NS1) variant (T363A) DNA.
XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
XX tumoural disease; gene therapy; mutant; mutein; variant; ds.
XX Parvovirus.
OS Synthetic.
XX
XX Key Location/Qualifiers
XX CDS 1..2019
XX FT /*tag= a
XX FT /product= "NS1 variant (T363A) protein"
XX FT replace(1187, A)
XX FT /*tag= b
XX
XX EP1077260-A1.
XX
XX 21-FEB-2001.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
```


QY 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACCAATATCTCTGTGGC 1860
DB 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACCAATATCTCTGTGGC 1860
QY 1861 GGCACCTGCAGAAACCCAGAACACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920
DB 1861 GGCACCTGCAGAAACCCAGAACACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920
QY 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTTCGGTGGCGGAACCG 1980
DB 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTTCGGTGGCGGAACCG 1980
QY 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACCTGGACTAA 2019
DB 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACCTGGACTAA 2019

RESULT 5

AAD02805
ID AAD02805 standard; DNA; 2019 BP.

XX AAD02805;

XX 06-AUG-2003 (revised)

DT 31-MAY-2001 (first entry)

XX Parvovirus non-structure protein 1 (NS1) variant (T463A) DNA.

DE NS1; non-structure protein 1; cytotstatic; gene therapy; toxin; therapy;

KW tumoral disease; gene therapy; mutant; mutein; variant; ds.

XX Parvovirus.

OS Synthetic.

XX Key

FT CDS

FT 1..2019

FT /*tag= a

FT /product= "NS1 variant (T463A) protein"

FT replace(1387, A)

FT /*tag= b

FT mutation

FT EP1077260-A1.

PN 21-FEB-2001.

XX 13-AUG-1999; 99EP-00115161.

XX 13-AUG-1999; 99EP-00115161.

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX Nueesch J; Rommelaere J;

XX WPI; 2001-212717/22.

XX P-PSDB; AAY72710.

XX Novel parvovirus non-structure protein variant, useful for treating

XX tumoral diseases, has a shifted equilibrium between DNA replication and

XX transcription activities, and cytotoxic activity.

XX Claim 7; Page 27-30; 4lpp; English.

XX The present sequence is a DNA encoding parvovirus non-structure protein 1

XX (NS1) variant (T463A). The invention relates to the variants of the

XX parvovirus non-structure protein (NS1) having a shifted equilibrium

XX between the DNA replication and transcription activities, and the

XX cytotoxicity activity. These variants are useful as toxins for treating

XX tumoral diseases. The variant DNAs are useful as vectors for gene

XX therapy. (Updated on 06-AUG-2003 to correct OS field.)

XX Sequence 2019 BP; 697 A; 389 C; 453 G; 480 T; 0 U; 0 Other;

XX Query Match 99.8%; Score 2015.8; DB 5; Length 2019;

Best Local Similarity 99.9%; Pred. No. 0;
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGCTCGGAAATCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTTAAAGGAAAA 60
DB 1 ATGCTCGGAAATCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTTAAAGGAAAA 60
QY 61 AGTAACGAGGAAGTGTCTCATTTGTTTAAAAATGAAATGTTCACTGAATGGAAAA 120
DB 61 AGTAACGAGGAAGTGTCTCATTTGTTTAAAAATGAAATGTTCACTGAATGGAAAA 120
QY 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTCGAGGAGCAGCTGGAATCTTTACAA 180
DB 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTCGAGGAGCAGCTGGAATCTTTACAA 180
QY 181 CGAGGAGCGGAAACTACTTTGGGACCAAGGAGCATGGAATGGGAAACACAGTGGAT 240
DB 181 CGAGGAGCGGAAACTACTTTGGGACCAAGGAGCATGGAATGGGAAACACAGTGGAT 240
QY 241 GAAATGACCAAAAGCAAGTATTTCACTTTTGGTAAATGTTTAAATGTTTATTTGAA 300
DB 241 GAAATGACCAAAAGCAAGTATTTCACTTTTGGTAAATGTTTAAATGTTTATTTGAA 300
QY 301 GTGCTTAACACAAAGAAATATTTCTCTGGTGTATTTAAATGTTTGGCAACATGAATGG 360
DB 301 GTGCTTAACACAAAGAAATATTTCTCTGGTGTATTTAAATGTTTGGCAACATGAATGG 360
QY 361 GGAAGAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGACTTTTGTCAAGCT 420
DB 361 GGAAGAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGACTTTTGTCAAGCT 420
QY 421 CAAGGAAATGTTGGAGAAAGGCAACTAAATGTTTACTGGAGCAGATGGTTGTAACAGCC 480
DB 421 CAAGGAAATGTTGGAGAAAGGCAACTAAATGTTTACTGGAGCAGATGGTTGTAACAGCC 480
QY 481 TGTAAATGTCNACTACACAGCTGAAGAAATTTAAACTAAGAAATAGCAGAGACAAT 540
DB 481 TGTAAATGTCNACTACACAGCTGAAGAAATTTAAACTAAGAAATAGCAGAGACAAT 540
QY 541 GAGTGGTTTACTCTTACTTACTTATAGCATAGCAAAACCAAAAGACTATACCAAGTGT 600
DB 541 GAGTGGTTTACTCTTACTTACTTATAGCATAGCAAAACCAAAAGACTATACCAAGTGT 600
QY 601 GTTCTTTTGGAAACATGATTTCTTATTTTAACTAAAAAGAAAAAAGACTATACCAAGTGT 660
DB 601 GTTCTTTTGGAAACATGATTTCTTATTTTAACTAAAAAGAAAAAAGACTATACCAAGTGT 660
QY 661 CCACCAAGAGACGAGGCTATTTCTTACAGTGTCTGGCTGGGAAAGCTTAACTTTTAA 720
DB 661 CCACCAAGAGACGAGGCTATTTCTTACAGTGTCTGGCTGGGAAAGCTTAACTTTTAA 720
QY 721 AAAGAAGGCGAGCGCCATCTAGTGAGCAAACTATACATGATGACATGCGGCCAGAAACG 780
DB 721 AAAGAAGGCGAGCGCCATCTAGTGAGCAAACTATACATGATGACATGCGGCCAGAAACG 780
QY 781 GTTGAACACACAGTAAACCACTGCGAGGAAACTAAGCGGCGGAGAAATCAAACTAAAAA 840
DB 781 GTTGAACACACAGTAAACCACTGCGAGGAAACTAAGCGGCGGAGAAATCAAACTAAAAA 840
QY 841 GAAGTTGCTTATTAACACTACATTAAGAGCTGGTGCATAAAGAGTAACTCACCAGAG 900
DB 841 GAAGTTGCTTATTAACACTACATTAAGAGCTGGTGCATAAAGAGTAACTCACCAGAG 900
QY 901 GACTGGATGATGATGAGCAGCAGTGTACATTGAAATGATGCTCAACAGCTGGAGAA 960
DB 901 GACTGGATGATGATGAGCAGCAGTGTACATTGAAATGATGCTCAACAGCTGGAGAA 960
QY 961 AACTCTCTGAAATAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAACCAACAGCA 1020
DB 961 AACTCTCTGAAATAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAACCAACAGCA 1020
QY 1021 TTTGACTTAAATTTTAAAGAAAGCTGAAACCCAGCAAACTAACCACTTTTCACTGCTGAC 1080
DB 1021 TTTGACTTAAATTTTAAAGAAAGCTGAAACCCAGCAAACTAACCACTTTTCACTGCTGAC 1080

```
Db 1021 TTTGACTTAATTTTAGAAAAAGCTGAACCCAGCAAACTAACCACTTTTCACTGCTGAC 1080
Qy 1081 ACAAGAACCTGCAGAAATTTTGTCTTTTATGCTGGAATCTATGTTAAAGTTTGGCAATGCT 1140
Db 1081 ACAAGAACCTGCAGAAATTTTGTCTTTTATGCTGGAATCTATGTTAAAGTTTGGCAATGCT 1140
Qy 1141 ATTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATCTGTTTATTTATCATGGACCA 1200
Db 1141 ATTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATCTGTTTATTTATCATGGACCA 1200
Qy 1201 GCAGACAGGCAAAATCTATTATGCAAGCCATAGCAAGCACTTGGCAATGTTGCT 1260
Db 1201 GCAGACAGGCAAAATCTATTATGCAAGCCATAGCAAGCACTTGGCAATGTTGCT 1260
Qy 1261 TGTATTAATGCAAGCAATGTAATCTTTTCAATTTAATGACTGTACCAACAGAACTTGATT 1320
Db 1261 TGTATTAATGCAAGCAATGTAATCTTTTCAATTTAATGACTGTACCAACAGAACTTGATT 1320
Qy 1321 TGGGTAGAAGAGCTGGTAACTTTTGGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT 1380
Db 1321 TGGGTAGAAGAGCTGGTAACTTTTGGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT 1380
Qy 1381 GGTCAAACTATTGCAATGTAATCAAAAGGAGGAGCAAGCAAGTGAACCAACCA 1440
Db 1381 GGTCAAACTATTGCAATGTAATCAAAAGGAGGAGCAAGCAAGTGAACCAACCA 1440
Qy 1441 GTCATATGACCAAAATGAGAACATTACAGTGGTCAAGATAGGCTGCGAAGAAAGACCA 1500
Db 1441 GTCATATGACCAAAATGAGAACATTACAGTGGTCAAGATAGGCTGCGAAGAAAGACCA 1500
Qy 1501 GAACACACTCAACCAATCAGACAGACAGATGCTTAAACATTAACATCAACATCTTGCT 1560
Db 1501 GAACACACTCAACCAATCAGACAGACAGATGCTTAAACATTAACATCAACATCTTGCT 1560
Qy 1561 GGTGACTTTGGTGTGACAAATGAATGCGCAATGTTGCTGTTGTTGTTAAAG 1620
Db 1561 GGTGACTTTGGTGTGACAAATGAATGCGCAATGTTGCTGTTGTTGTTAAAG 1620
Qy 1621 AATGTTTACCAATCTACATGGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCTGATTGG 1680
Db 1621 AATGTTTACCAATCTACATGGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCTGATTGG 1680
Qy 1681 TCAGAAACTGGGCGAGGCAAGGTCGCAACTCTTAAATTAATTAATTAAGTTGCGCACGC 1740
Db 1681 TCAGAAACTGGGCGAGGCAAGGTCGCAACTCTTAAATTAATTAATTAAGTTGCGCACGC 1740
Qy 1741 TCACCATTCAGCACACCGAAAGTACGCTCTCAGCCAGAACATGCACTAATCTCACTT 1800
Db 1741 TCACCATTCAGCACACCGAAAGTACGCTCTCAGCCAGAACATGCACTAATCTCACTT 1800
Qy 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAATATCTCTGTTGG 1860
Db 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAATATCTCTGTTGG 1860
Qy 1861 GGCACTCGAGAAACCCAGAACACTGGGGAAAGCTGGTTCCAAAGCCTGCCAAGATGTCAA 1920
Db 1861 GGCACTCGAGAAACCCAGAACACTGGGGAAAGCTGGTTCCAAAGCCTGCCAAGATGTCAA 1920
Qy 1921 CTGAGCCCAACTGGTTCAGAGATGAGGAGGATTTGAGAGCGTCTTCGTTGCGGAACCG 1980
Db 1921 CTGAGCCCAACTGGTTCAGAGATGAGGAGGATTTGAGAGCGTCTTCGTTGCGGAACCG 1980
Qy 1981 TTGAAGAAAGACTTTCAGCGAGCCGCTGAACCTGGACTAA 2019
Db 1981 TTGAAGAAAGACTTTCAGCGAGCCGCTGAACCTGGACTAA 2019
```

RESULT 6

AAD02803

ID AAD02803 standard; DNA; 2019 BP.

XX

AC AAD02803;

XX

```
DT 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
DE Parvovirus non-structure protein 1 (NS1) variant (T394A) DNA.
KW NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
tumoural disease; gene therapy; mutant; mutein; variant; ds.
OS Parvovirus.
OS Synthetic.
XX Location/Qualifiers
FH 1..2019
FT /tag= a
FT /product= "NS1 variant (T394A) protein"
FT mutation replace(1180, A)
FT /tag= b
XX EP1077260-A1.
XX 21-FEB-2001.
XX 13-AUG-1999; 99EP-00115161.
XX 13-AUG-1999; 99EP-00115161.
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX Nueesch J, Rommelaere J;
XX WPI; 2001-212717/22.
XX P-PSDB; AAY72708.
XX Novel parvovirus non-structure protein variant, useful for treating
tumoral diseases, has a shifted equilibrium between DNA replication and
transcription activities, and cytotoxic activity.
XX Claim 7; Page 22-24; 4lpp; English.
XX The present sequence is a DNA encoding parvovirus non-structure protein 1
(NS1) variant (T394A). The invention relates to the variants of the
parvovirus non-structure protein (NS1) having a shifted equilibrium
between the DNA replication and transcription activities, and the
cytotoxicity activity. These variants are useful as toxins for treating
tumoral diseases. The variant DNAs are useful as vectors for gene
therapy. (Updated on 06-AUG-2003 to correct OS field.)
XX Sequence 2019 BP; 697 A; 389 C; 453 G; 480 T; 0 U; 0 Other;
Query Match 99.8%; Score 2015.8; DB 5; Length 2019;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAAA 60
Db 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAAA 60
Qy 61 AGTAACCCAGGAAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGGAATA 120
Db 61 AGTAACCCAGGAAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGGAATA 120
Qy 121 GATATCGGATGGATAGTTTACAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTACAA 180
Db 121 GATATCGGATGGATAGTTTACAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTACAA 180
Qy 181 CGAGGAGCGGAATACTACTTGGGACCAAAAGCAGGACATGGAATGGGAAACCCAGTGGAT 240
Db 181 CGAGGAGCGGAATACTACTTGGGACCAAAAGCAGGACATGGAATGGGAAACCCAGTGGAT 240
Qy 241 GAAATGACCAAAAAAGCAAGATTTTCATTTTGGTTTAAATGTTTATTGAA 300
Db 241 GAAATGACCAAAAAAGCAAGATTTTCATTTTGGTTTAAATGTTTATTGAA 300
```

301 GTGCTTAACACAAAGAAATATATTTCCCTGGTGATGTTAAATGGTTTGTGCAACATGATGG 360
301 GTGCTTAACACAAAGAAATATATTTCCCTGGTGATGTTAAATGGTTTGTGCAACATGATGG 360
361 GGAAGACACCAAGCTCGCACTGCATGTACTAATTCGAGGAAGAGACTTTAGTCAAGCT 420
361 GGAAGACACCAAGCTCGCACTGCATGTACTAATTCGAGGAAGAGACTTTAGTCAAGCT 420
421 CAAGGGAATGGTGGAGGAAGCAACTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCC 480
421 CAAGGGAATGGTGGAGGAAGCAACTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCC 480
481 TGTAAATGTGCAACTAACACACAGCTGAAAGAAATTAACCTAGAGAGAAATAGCAGAGCAAT 540
481 TGTAAATGTGCAACTAACACACAGCTGAAAGAAATTAACCTAGAGAGAAATAGCAGAGCAAT 540
541 GAGTGGGTACTCTACTTACTTAAAGCATAAGCAACCAAAAGAAAGACTATACCAAGTGT 600
541 GAGTGGGTACTCTACTTACTTAAAGCATAAGCAACCAAAAGAAAGACTATACCAAGTGT 600
601 GTTCTTTTGGAAACATGATGTTGCTTACTATTTTAACTAAAGAAAGAAATAGCAGCTAGT 660
601 GTTCTTTTGGAAACATGATGTTGCTTACTATTTTAACTAAAGAAAGAAATAGCAGCTAGT 660
661 CCACCAAGACGAGGAGCTATTTCTTAGCAGTGACTCTGGCTGGAAAACTTAACCTTTTA 720
661 CCACCAAGACGAGGAGCTATTTCTTAGCAGTGACTCTGGCTGGAAAACTTAACCTTTTA 720
721 AAAGAGCGGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG 780
721 AAAGAGCGGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG 780
781 GTTGAAACCAAGTAACTACACTCGCGCAGAACTAAAGCGCGCAGAAATTCAAACTAAAAA 840
781 GTTGAAACCAAGTAACTACACTCGCGCAGAACTAAAGCGCGCAGAAATTCAAACTAAAAA 840
841 GAAGTGTCTATTAAACTACACTTAAAGAGCTGGTGATTAAGAGTAACCTACACAGAG 900
841 GAAGTGTCTATTAAACTACACTTAAAGAGCTGGTGATTAAGAGTAACCTACACAGAG 900
901 GACTGGATGATGACGAGCAGACAGTTTACATTGAATGATGCTCAACACGAGTGGAGAA 960
901 GACTGGATGATGACGAGCAGAGTTTACATTGAATGATGCTCAACACGAGTGGAGAA 960
961 AACCTGCTGAAAAATACGCTAGAGATTGTACACTAATCTAGCCAGAACCAAAACAGCA 1020
961 AACCTGCTGAAAAATACGCTAGAGATTGTACACTAATCTAGCCAGAACCAAAACAGCA 1020
1021 TTTGACTTAATTTTAGAAAAAGCTGAAACCGAGCAAACTAACCAACTTTTCACTGCCTGAC 1080
1021 TTTGACTTAATTTTAGAAAAAGCTGAAACCGAGCAAACTAACCAACTTTTCACTGCCTGAC 1080
1081 ACAAGACCTGCAGAAATTTTGTCTTTTTCATGGCTGGAACTATGTTAAAGTTTGCATGCT 1140
1081 ACAAGACCTGCAGAAATTTTGTCTTTTTCATGGCTGGAACTATGTTAAAGTTTGCATGCT 1140
1141 ATTTGCTGTGTTTAAACAGACAGGAGCAAAAGAAATATCTGTTTATTTTATTCATGGACCA 1200
1141 ATTTGCTGTGTTTAAACAGACAGGAGCAAAAGAAATGCTGTTTATTTTATTCATGGACCA 1200
1201 GCCAGCAGGCAAACTATATTATTCACAAAGCCATAGCAACAGCAGTTGGCAATGTTGGT 1260
1201 GCCAGCAGGCAAACTATATTATTCACAAAGCCATAGCAACAGCAGTTGGCAATGTTGGT 1260
1261 TGCTATTAATCGAGCAATGTAATTTTCCATTTTAACTGATGCTACCAACAGAACTTGATT 1320
1261 TGCTATTAATCGAGCAATGTAATTTTCCATTTTAACTGATGCTACCAACAGAACTTGATT 1320
1321 TGGGTAGAAGAGCTGGTAACTTTGGACAGCAAGTAAACAGTTTAAAGCCATTGCTCT 1380
1321 TGGGTAGAAGAGCTGGTAACTTTGGACAGCAAGTAAACAGTTTAAAGCCATTGCTCT 1380
1381 GGTCAAACTATTGCAATTGATCAAAAAGGAAAGGCAAGCAACAGATTTGAACCAACACCA 1440

1381 GGTCAAACTATTGCAATTGATCAAAAAGGAAAGGCAAGCAACAGATTTGAACCAACACCA 1440
1441 GTCATCATGACCAAAATGAGAACATTACAGTGGTCAGAAATAGGCTCGGAAGAAAGACCA 1500
1441 GTCATCATGACCAAAATGAGAACATTACAGTGGTCAGAAATAGGCTCGGAAGAAAGACCA 1500
1501 GAACACACTCAACCAATTCAGAGACAGAAATGCTTTAACTTCACTTAAACACATACCTTGCCT 1560
1501 GAACACACTCAACCAATTCAGAGACAGAAATGCTTTAACTTCACTTAAACACATACCTTGCCT 1560
1561 GGTGACTTTGGTTTGGTTGACAAAATGAATGGCCCATGATTTGTGCTTGGTTGGTAAAG 1620
1561 GGTGACTTTGGTTTGGTTGACAAAATGAATGGCCCATGATTTGTGCTTGGTTGGTAAAG 1620
1621 AATGGTTTACCAATCTACCATGGCAAGCTACTGTGCTCTAAATGGGGCAAAAGTTCTGTATTGG 1680
1621 AATGGTTTACCAATCTACCATGGCAAGCTACTGTGCTCTAAATGGGGCAAAAGTTCTGTATTGG 1680
1681 TCAGAAAACTGGGGGAGGCCAAAGGTGCGCAACTCTCTATAAAATTTACTAGGTTGCGGACGC 1740
1681 TCAGAAAACTGGGGGAGGCCAAAGGTGCGCAACTCTCTATAAAATTTACTAGGTTGCGGACGC 1740
1741 TCACCAATTCACGACACCGAAAAGTACGCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800
1741 TCACCAATTCACGACACCGAAAAGTACGCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800
1801 GCATCGGATCTCAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAAAATCTCTCTGTTGCG 1860
1801 GCATCGGATCTCAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAAAATCTCTCTGTTGCG 1860
1861 GGCACTGCGAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920
1861 GGCACTGCGAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920
1921 CTGAGCCCACTTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCCGTTGCGGAACCG 1980
1921 CTGAGCCCACTTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCCGTTGCGGAACCG 1980
1981 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019
1981 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019

RESULT 7

AAT15311

ID AAT15311 standard; DNA; 5049 BP.

XX AAT15311;

XX DT 14-OCT-1996 (first entry)

XX Non-attenuated canine parvovirus CPV-39 passage 5 DNA.

XX Parvovirus; dog; vaccine; CPV; myocardial disease; enteric disease; ds;

XX ss.

XX Canine parvovirus.

XX OS

XX PN W09614088-A1.

XX PD 17-MAY-1996.

XX PF 02-NOV-1995; 95WO-US014207.

XX PR 08-NOV-1994; 94US-0036345.

XX PA (CORR) CORNELL RES FOUND INC.

XX PI Parriah CR, Gruenberg A, Carmichael LE;

XX XX WPI; 1996-251556/25.

XX

QY 1915 GGTCACTGAGCCCACTTGGTCAGAGATCGAGAGGATTGAGAGCGTCTCGGTGCG 1974
Db 2181 GTCAAGCGAGTCGCGTGTGTCGAAATAGAGCAGACCTGAGAGCCATCTTTACTTCT 2240
QY 1975 GAACCGTTGAAGAAGACTTCAGCGA 2000
Db 2241 GAACAAITGGAGAAGATTTCGAGA 2266

RESULT 8

AAT15312

ID AAT15312 standard; DNA; 5049 BP.

XX AC AAT15312;

XX XX 14-OCT-1996 (first entry)

XX DE Attenuated canine parvovirus CPV-39 passage 60 DNA.

XX XX Parvovirus; dog; vaccine; CPV; myocardial disease; enteric disease; ds;

KW KW ss.

XX OS Canine parvovirus.

XX FH Key Location/Qualifiers

FT misc_feature 59

FT FT /*tag= a

FT FT /note= "A, C or T"

FT FT misc_feature 97

FT FT /*tag= b

FT FT /note= "A, G or T"

FT FT misc_feature 4745

FT FT /*tag= c

FT FT /note= "A G or C"

FT FT misc_feature 4881

FT FT /*tag= d

FT FT /note= "A G or T"

XX XX WO9614088-A1.

XX XX 17-MAY-1996.

XX PF 02-NOV-1995; 95WO-US014207.

XX PR 08-NOV-1994; 94US-00336345.

XX XX (CORR) CORNELL RES FOUND.INC.

XX PA Parrish CR, Gruenberg A, Carmichael LE;

XX XX WPI; 1996-251556/25.

XX DR Attenuated CPV strains contg. up to 4 mutation (s) relative to control

XX PT virus - useful as a veterinary vaccine against CPV disease in animals,

XX PT such as wild or domestic dogs.

XX XX Claim 2; Page 24-27; 42pp; English.

XX XX This viral DNA is isolated from an attenuated CPV. The DNA is preferably

XX CC derived from vB140. The DNA is cloned into a vector which is used to

XX CC transfect a host cell. The vector used is preferably pGEM32 or pGEM52.

XX CC The host cells to be transfected are selected from Norden Laboratory

XX CC feline kidney cells, mink lung cells, Madin-Darby canine kidney cells or

XX CC canine A72 cells

XX XX Sequence 5049 BP; 1790 A; 814 C; 1029 G; 1416 T; 0 U; 0 Other;

XX XX Query Match 51.5%; Score 1040.4; DB 2; Length 5049;

XX XX Best Local Similarity 71.4%; Pred. No. 4.6e-268;

XX XX Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;

QY 1 ATGCGTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 60

Db 273 ATGTCGGCAACAGTATATCTGAGGAAGTTATGAGGGAGTAAATTGGTTAAAGAAACAT 332
QY 61 AGTAAACGAGGAAGTGTCTCATTTGTTTTTAAAAATGAAAATGTTTCAACTGAATGGAAAA 120
Db 333 GCAGAAAAATGAAGCAATTTTCGTTTTGTTTTTAAATGTGACAACGTCCTCAACTAAATGGAAG 392
QY 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTCGAGGAGCAGAGCTGAAATCTTTTACAA 180
Db 393 GATGTTTCGCTGGAAACAACTATACCAAAACCAATTTCAAAATGAAGAGCTAACTCTTTAAAT 452
QY 181 CGAGGAGCGAAACTACTCTGGGACCAAGC-----GAGGACATGGAATGGGAACACCA 234
Db 453 AGAGGAGCACAACAGCAATGGATCAACCGAAGAGAAGAAATGGACTGGGAATCGGAA 512
QY 235 GTGGATGAAATGACCAAAAAAGCAAGTATTCATTTTGTGATCTTTTGGTTAAAAAATGTTTA 294
Db 513 GTTGATAGTCTCGCCAAAAAGCAAGTACAAACTTTTGTGATCAATTAATTAATAAATGCTT 572
QY 295 TTTGAAGTGTGTTAACACAAGAAATATATTTCTGGTGATGTTTAATTTGGTTTGGCAACAT 354
Db 573 TTTGAAGTCTTTGTTCTAAAAATATAGAACCAAAATGAATGTGTGTTGTTTATTTCAACAT 632
QY 355 GAATGGGAAAGACCAAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 414
Db 633 GAATGGGAAAGAAATCAAGGCTGGCATTTGTCATGTTTTTCTTCTATAGTAAGAACTTACAA 692
QY 415 CAAGCTCAAGGAAATGGTGGAGAAGGCAACTAAATGTTTTTACTGGAGCAGATGTTGGTA 474
Db 693 CAAGCAACTGTTAAATGGCTACGACACAATGAATATGTTATTTGGAGTAGATGTTGGTG 752
QY 475 ACAGCCTGTAATGTGCAACTAACACAGCTGAAAGAAATTTAACTAAGAGAAATAGCAGAA 534
Db 753 ACTCTTTGTTGCGTAAACTTTAACCAAACTGAAAGAAATTTAAGCTCAGAGAAATTCAGAA 812
QY 535 GACAATGAGTGGGTACTCTCTACTTACTTATAAGCATAAGCAACCAACCAAGAAAGACTATACC 594
Db 813 GATAGTAATGGGTGACTATATTAAACATACAGACATAAGCAACCAACCAAGAAAGACTATGTT 872
QY 595 AAGTGTGTTCTTTTGGAAACATGATTCCTTACTATTATTTTAACTAAAAAGAAATTAAGC 654
Db 873 AAAATGTTTCATTTTGGAAATATGATAGCATATTACTTTTAAACAAAGAAATTAATGTC 932
QY 655 ACTAGTCCCAAGAGAGAGAGGAGCTATTTCTTTAGCAGTGAAGTCTCTGCTGGAAGAACTAAC 714
Db 933 CACAT---GACAAAAAGAAAGTGGCTATTTTAAAGTACTGATTTCTGCTGGAATTTTAAAC 989
QY 715 TTTTAAAGAAAGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGATGATGATGCGGCCA 774
Db 990 TTTATGAAGTATCAAGACAGACAAAATTTGTGAGCACACTTTTACACTGAAACAAATGAACCA 1049
QY 775 GAAACGTTGAAACACAGTAACCACTGCGCAGGAAACTAAGCGCGGAGAAATTTCAAACT 834
Db 1050 GAAACGTTGAAACACAGTGTGACAGACAGCAAGCAAGCAAGCGCGGAGAAATTTCAAACT 1109
QY 835 AAAAAAGAAAGTGTGCTATTAATAACTACACTTAAAGAGCTGGTGCATTAAGAGTAACCTCA 894
Db 1110 AAAAAAGAAAGTGTCAATCAATGTAATCTTCTGCGGACTTGGTTAGTAAAGAGTAACATCA 1169
QY 895 CCAGAGGACTGGATGATGATGAGCCAGACAGTGTACATTCGAAATGATGGCTCAACACAGGT 954
Db 1170 CTTGAAGACTGGATGATGTTTACAAACAGATAGTTATTTGAAATGATGGCAGCAACACAGGA 1229
QY 955 CGAGAAAACTGCTGAAAAAATACGCTAGAGATTTGTACACTAACTCTTAGCAGAACCAAAA 1014
Db 1230 GGTGAAATCTTTTAAAAAATACACTTGAATTTTGTACTTTGACTTTAGCAGAACCAAAA 1289
QY 1015 ACAGCAATTTGACTTAATTTTAAAAAAGCTGAAACCAAGCAAACTAACCAACTTTTCACTG 1074
Db 1290 ACAGCAATTTGAAATTAATTAATCTTGAAGAAAGCAGATAACTAACTAACTAACTTTGATCTT 1349
QY 1075 CTTGACACAAGAACTGACAGAAATTTTGTCTTTTCTGCTGGAACACTATGTTAAAGTTGC 1134

Db 1350 GCAAAATTTAGAACATGTCAAATTTTGAATGACGAGTGAATGGATTAAGATTTGT 1409
Qy 1135 CATGCTATTGCTGCTTTTAAACACAGCAAGGAGGCAAAAGAAATCTGTTTATTTCAT 1194
Db 1410 CAGCCTATAGCATGTGTTTAAATAGACAGGTGGTAAAGAAATACAGTCTTTTCAT 1469
Qy 1195 GGACCAAGCCAGCAGCGCAAAATCTATTATTGCAAGCCATAGCACAAGCTGTGGCAAT 1254
Db 1470 GGACCAAGCAAGTACAGGAAATCTATCAITGCTCAAGCCATAGCACAAGCTGTGGTAA 1529
Qy 1255 GTTGGTTGCTATTAATGAGCAAGCAATTAACCTTTCCATTTTAAGTACGTACCAACAGAC 1314
Db 1530 GTTGGTTGTTATTAATGAGCAAGCAATTTTCCATTTTAAGTACGTACCAATAAAAAT 1589
Qy 1315 TTCAATTTGGGTAGAGAGCTGTAATTTTGGACAGCAAGTAAACAGTTTAAAGCAAT 1374
Db 1590 TTAATTTGGATTGAAGAGCTGTAATTTTGGTCAACAGTTTAAATTTAAAGCAATC 1649
Qy 1375 TGCTCTGGTCAAACTATTTCGCAATTTGATCAAAAAGGAAAGGCAAGCAAGATTTAAAGCAAT 1434
Db 1650 TGTTCTGGACAAACAAATTGAATTTGATCAAAAAGGTAAGCAAAATTTGAACCA 1709
Qy 1435 ACACGATCATATGACCAACAAATAGAACATTTACAGTGTGATAGTACGCTGCGAAGAA 1494
Db 1710 ACTCCAGTAAATTATGACAACTAATGAATATAACAAATTTGTGAAGATTTGATGTGAAGAA 1769
Qy 1495 AGACCAAGACACACTCAACCAATCAGAGACAGAAATGCTTAAACATTTCACTAAACATACC 1554
Db 1770 AGACCTGACATACACAAACCAATAGAGACAGAAATGTTGACATTTAAGTTAGTATGAAG 1829
Qy 1555 TTGCTCTGGTGAATTTGGTTGGTGAACAAAATGAATGGCCCATGATTTGTGCTGGTTG 1614
Db 1830 CTTCAGGAGACTTTGGTTGGTTGATAAAGAAAGATGGCTTTAATATGTGATGCTGTTA 1889
Qy 1615 GTAAAGAAATGTTACCAATCTACATGGCAAGTCTGCTAAATGGGGCAAGTTTCT 1674
Db 1890 GTTAAACATGTTTGAATCAACCATGGCTAACTATACATCATTTGGGGAAAGTACCA 1949
Qy 1675 GATTGGTTCAGAAACTGGGGGAGCCAAAGTGCCAACTCCTATATAATTTACTAGGTTCG 1734
Db 1950 GAATGGATGAAACTGGGGGAGCCTTAATACAGAAAGTATAAATTCACCAGGTTCG 2009
Qy 1735 GCACGCTCACCAATTCACGACACCGAAAGTACGCTCTCAGCAGCAAGTATGACACT 1794
Db 2010 AAAGACT---TAGAGACACAGCGCAAGCAATCCTCAGAGTCAAGACCAAGTTCTAACT 2066
Qy 1795 CCATTCGATCGATCTCGAGGACCTGCTTTAGAGCCTTTGGAGCACACCAAAATCTCCT 1854
Db 2067 CCTCTGACTCCGAGCGTAGTGGACCTTTGCACTGGAACCGTGGAGTACTCCAGATACGCT 2126
Qy 1855 GTTCGGGCACTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGAT 1914
Db 2127 ATTGCAGAACTGCAATCAACATCAACCAACTTGGCGTTACTCA-----CAAAGAC 2180
Qy 1915 GGTCAACTGACCCCAACTTGTGTGATGATGAGAGGATTTGAGAGCGTCTCGGTGCG 1974
Db 2181 GTGCAAGCGATCGACGCTGCTGCGGAAATAGAGGACAGCCTGAGAGCCATCTTCTTCT 2240
Qy 1975 GAACCGTTGAAGAAGACTTTCAGCGA 2000
Db 2241 GAACAAATTGAAGAAGATTTTCGAGA 2266

RESULT 9

ID AAT88324 standard; DNA; 5049 BP.

XX AC AAT88324;

DT 17-OCT-2003 (revised)

XX DT 21-MAY-1998 (first entry)

DE Attenuated canine parvovirus (vBI440) genomic DNA.

XX Canine parvovirus; CPV; attenuation; vBI440; vaccine; dog;
KW feline panleukopenia virus; mink enteritis virus; infection; ds.
XX Canine parvovirus; vBI440 (ATCC VR 2489).

Key Location/Qualifiers

FT mutation 59
FT /tag= c
FT /note= "base 59 is G in CPV-39 (passage 5)"
FT mutation 97
FT /tag= d
FT /note= "base 97 is C or T in CPV-39 (passage 5)"
FT CDS 273..2279
FT /tag= a
FT /note= "NS1/NS2 coding region"
FT CDS 2286..4541
FT /tag= b
FT /note= "VP1/VP2 coding region"
FT mutation 4745
FT /tag= e
FT /note= "base 4745 is T in CPV-39 (passage 5)"
FT mutation 4881
FT /tag= f
FT /note= "base 4881 is C in CPV-39 (passage 5)"

XX WO9742972-A1.

XX 20-NOV-1997.

XX 06-MAY-1997; 97WO-US007584.

XX 15-MAY-1996; 96US-00647655.

XX (CORR) CORNELL RES FOUND INC.

XX Parrish CR, Carmichael LE, Gruenberg A;

XX WPI; 1998-008583/01.

XX Canine parvovirus DNA carrying specific attenuating mutation(s) - used as
PT vaccines for protection against parvovirus and feline pan-leukopenia
XX virus infections.

PS Example 8; Page 34-37; 60pp; English.

XX This DNA sequence comprises an attenuated virus genome derived by serial
CC passaging (60 times) of virulent canine parvovirus (CPV) type 2b isolate
CC 39 in NL/K feline kidney host cells. The attenuated virus is designated
CC vBI440 (ATCC VR 2489). It contains 4 mutations relative to the sequence
CC (see AAT88320) of the control (5th passage) wild-type CPV-2b. 2 Mutations
CC are within the hairpin formed by the 3' terminal palindromic: the mutation
CC at nucleotide 59 introduces an A into a G-C rich region within the tip of
CC the hairpin, disrupting the base pairing in one of the 2 small internal
CC palindromes within that sequence; the thymine at nucleotide 97 is
CC adjacent to the mismatched bubble (flip-flop) sequence within the
CC palindromic. The DNA from attenuated CPV strains (see also AAT88321) is
CC used for the production of infectious molecular DNA clones, which, in
CC turn, can be transfected into cells to generate master stocks of the
CC virus. The attenuated viruses can be used in dogs as a vaccine to protect
CC against CPV disease, or more generally in cats and minks to protect
CC against feline panleukopenia virus and mink enteritis virus. The vaccines
CC protect against the currently prevalent CPV-2b type (and all extant
CC strains of types 2 and 2a), providing a long term immune response.
CC (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 5049 BP; 1791 A; 812 C; 1029 G; 1417 T; 0 U; 0 Other;

XX Query Match 51.5%; Score 1040.4; DB 2; Length 5049;

XX Best Local Similarity 71.4%; Pred. No. 4.6e-268;

XX Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;

Qy 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGTTAAAGGAAAA 60


```
Db 1230 GGTGAAATCTTTAAAAAATACACTTGAAATTTGTACTTTGACTTTTAGCAAGAACAAAA 1289
Qy 1015 ACAGATTGACTTAATTTTAGAAAAAGCTGAAACACGACAAACTAAACAACTTTTCACTG 1074
Db 1290 ACAGCAATTGAATTAATCTTTGAAAAAGCAGATAATACTAAACTAACTAACTTTGATCTT 1349
Qy 1075 CCTGACACAGAACTCGAGAAATTTTTCCTTTTCATCGCTGGAATATGTTAAAGTTTGC 1134
Db 1350 GCAATTTCTAGACATGTCAAATTTTGTAGATGCACGGATGGAATTTGGATTTAAAGTTTGT 1409
Qy 1135 CATCTATTGCTGTGTTTTTAAACAGACAGAGGCAAAAGAAATACCTGTTTTTATTTTCAT 1194
Db 1410 CACGCTATGATGTTGTTTTTAAATAGACAAAGTGTAAAGAAATACAGTTCTTTTTCAT 1469
Qy 1195 GGACGACGACAGCAGCAATCTATTATTGCACAAGCCATAGCAAGCAGTTGGCAAT 1254
Db 1470 GGACGACGACATGACAGAAATCTATCATTTGCTCAAGCCATAGCAAGCTGTGGGTAAT 1529
Qy 1255 GTTGCTGCTATATGATGAGCAATGTAATCTTTCCATTTAATGACTGTACCAAGAAC 1314
Db 1530 GTTGCTGCTATATGATGAGCAATGTAATCTTTCCATTTAATGACTGTACCAATAAAT 1589
Qy 1315 TTGATTTGGGTAGAAAGCTGGTAACTTTTGACAGCAAGTAAACCACTTTTAAAGCCATT 1374
Db 1590 TTAATTTGGATGGAAGCTGGTAACTTTGGTCAACAGTTAATCAATTTAAAGCAATC 1649
Qy 1375 TGCTCTGGTCAAACTATTGCAATTTGATCAAAAAGGAAAGGACAGCAACAGATTGAACCA 1434
Db 1650 TGTCTGGACAAACAAATTAGAAATTTGATCAAAAAGGTAAGCAAAATTTGAACCA 1709
Qy 1435 ACACCACTCATGACCACAAAATGAGAACATTTACAGTGTGTCAGAAATAGGCTGGGAAGAA 1494
Db 1710 ACTCCAGTAATATTGACAACTAATGAAATTTTAACTTTGTAAGATTTGATGTAAGAA 1769
Qy 1495 AGACCAAGACACACTCAACCAATCAGAGACAGAAATGCTTAACTTTTCACTAAACATACACC 1554
Db 1770 AGACCTGACATACACAAACCAATAGAGACAGAAATGTTGACATTAATGATGTAAG 1829
Qy 1555 TTGCTCTGGTGAATTTGGTTGGTTGACAAAAAATGAATGGCCCATGATTTGTGCTGGTTG 1614
Db 1830 CTTCCAGGAGACTTTGGTTGGTTGTAAGAAAGAAATGGCTTTTAAATATGTCATGTTA 1889
Qy 1615 GTAAAGAAATGTTTACCAATCTACATGCGCAAGCTACTGTGCTAAATGGGCAAAAGTTCTT 1674
Db 1890 GTTAAACATGTTTGTGAATCAACCATGCTAACTATACATCATCTTGGGAAAAAGTACCA 1949
Qy 1675 GATTGGTCAGAAACTGGGCGAGCCAAAGGTGCCAACTCTCTATAAATTTTACTAGGTTCG 1734
Db 1950 GAATGGATGAAACTGGGCGAGCCCTAAATACAGAGGTATAAATTCACCGAGTTGC 2009
Qy 1735 GCAGCTCACCATTTCAGCACACGAAAGTACGCTCTCAGCCAGAACTATGCACCTAACT 1794
Db 2010 AAAGACT---TAGAGACACAAGCGGCAAGCAATCTCAGAGTCAAGACCAAGTTCTAACT 2066
Qy 1795 CCATTGCTCGATCTCGAGGACCTGCTTTAGAGCCCTTGGAGCACACCAATATCTCT 1854
Db 2067 CCTCTGACTCGGACGCTAGTGGACCTTGGCACTGGCAACCGTGGAGTACTCCAGATACGCT 2126
Qy 1855 GTTGGCGGCACTGCAGAAACCCAGAACACTGGGGAAGCTGTTCCAAAGCTGCCAAGAT 1914
Db 2127 ATTGCAGAACTGCAATTAACAACTCAACAACTTGGCGTTACTCA-----CAAAGAC 2180
Qy 1915 GGTCAACTGAGCCCAACTTGGTCAGAGATCGAGGAGATTTGAGAGCGTCTCGGTGCG 1974
Db 2181 GTGCAAGGAGTCCGAGCTGCTCGAATATAGAGGACACCTGAGAGCCATCTTTACTTCT 2240
Qy 1975 GAACCGTTGAAGAAGACTTCAGCGA 2000
Db 2241 GAACAATTTGGAAGAAGATTTTCGAGA 2266
```

RESULT 11
AAT88320

```
ID AAT88320 standard; DNA; 5049 BP.
XX AC
XX AAT88320;
XX AC
DT 17-OCT-2003 (revised)
DT 21-MAY-1998 (first entry)
XX XX
DE Canine parvovirus 39 passage #5 (wild-type).
XX Canine parvovirus; CPV; attenuation; vB1440; vaccine; dog;
XX feline panleukopenia virus; mink enteritis virus; infection; ds.
XX Canine parvovirus; type 2b isolate 39.
XX
FH Key Location/Qualifiers
CDS 273..2279
FT /*tag= a
FT /note= "NS1/NS2 coding region"
FT CDS 2286..4541
FT /*tag= b
FT /note= "VP1/VP2 coding region"
XX
XX WO9742972-A1.
XX
XX 20-NOV-1997.
XX
XX 06-MAY-1997; 97MO-US007584.
XX
XX 15-MAY-1996; 96US-00647655.
XX
XX (CORR ) CORNELL RES FOUND INC.
XX
XX Parrish CR, Carmichael LE, Gruenberg A;
XX WPI; 1998-008583/01.
XX
XX Canine parvovirus DNA carrying specific attenuating mutation(s) - used as
XX vaccines for protection against parvovirus and feline pan-leukopenia
XX virus infections.
XX
XX Example 8; Page 37-40; 60pp; English.
XX
XX This DNA sequence comprises the genome of virulent canine parvovirus type
XX 2b (CPV-2b) isolate 39, obtained after 5 serial passages in NUKF feline
XX kidney host cells. Further passaging has yielded attenuated virus vB1440
XX (ATCC VR 2489) (see AAT88324) from the 60th passage, and a claimed virus
XX (see AAT88321) derived from the 65th passage (ATCC 2528). These
XX respectively contain 4 and 6 mutations in comparison to the virulent 5th
XX passage virus. The DNA from attenuated CPV-2b strains is used for the
XX production of infectious molecular DNA clones, which, in turn, can be
XX transfected into cells to generate master stocks of the virus. The
XX attenuated viruses can be used in dogs as a vaccine to protect against
XX CPV disease, or more generally in cats and minks to protect against
XX feline panleukopenia virus and mink enteritis virus. The vaccines protect
XX against the currently prevalent CPV-2b type (and all extant strains of
XX types 2 and 2a), and provide a long term immune response. (Updated on 17-
XX OCT-2003 to standardise OS field)
XX
SQ Sequence 5049 BP; 1789 A; 813 C; 1030 G; 1417 T; 0 U; 0 Other;
Query Match 51.5%; Score 1038.8; DB 2; Length 5049;
Best Local Similarity 71.3%; Pred. No. 1.2e-267;
Matches 1431; Conservative 0; Mismatches 557; Indels 18; Gaps 4;
Qy 1 ATGCGTCGGAATGCTTACTCTGATGAAGTTTTCGGAGCAACCACTGTTTAAAGGAAAA 60
Db 273 ATGTCGGCAACCACTATATCTGAGGAGTTATGAGGAGTAATTTGTTTAAAGAACAT 332
Qy 61 AGTAAACAGGAAGTGTCTCATTTGTTTGTAAAAATGAAATGTTCAACTGAATGAAAA 120
Db 333 GCAGAAAAATGAAGCATTTTCGTTGTTTAAATGTGACACACGCTCAACTAAATGGAAG 392
Qy 121 GATATCGGATGGAATAGTTACAAAAAAGAGCTGCAGAGGACGAGCTGTAATCTTTACAA 180
```



```

FT      /*tag= a
FT      /note= "see AAP40306"
FT      CDS      2107..3522
FT      /*tag= b
FT      /note= "see AAP40675"
XX
XX      WO8402847-A.
XX
XX      02-AUG-1984.
XX
XX      19-JAN-1984;      84WO-US000063.
XX
XX      19-JAN-1983;      83US-00459203.
XX      06-JAN-1984;      84US-00567968.
XX
XX      (AMGE-) AMGEN.
XX
XX      Fox GM;
XX
XX      WPI; 1984-201354/32.
XX      P-PSDB; AAP40306, AAP40675.
XX
XX      Polypeptide obtd. by recombinant DNA methods - for vaccination against
XX      parvovirus infections in man and animals.
XX
XX      Claim 10; Table II, Page 33-49; 80pp; English.
XX
XX      The inventors claim an immunologically active polypeptide for the
XX      development of vaccinal immunity against parvovirus infection. Also
XX      claimed are DNA sequences wholly or partly duplicative of defined
XX      sequences. The polypeptides are used in vaccines for conferring
XX      protection against parvovirus infections in man and animals. (Updated on
XX      24-OCT-2003 to standardise OS field)
XX
XX      SQ      Sequence 3524 BP; 1400 A; 719 C; 648 G; 757 T; 0 U; 0 Other;

Query Match      44.7%; Score 902.6; DB 1; Length 3524;
Best Local Similarity 69.8%; Pred. No. 3.6e-231;
Matches 1297; Conservative 0; Mismatches 544; Indels 18; Gaps 5;

QY      4 GCTGGAAATGCTTACTCTGATGAAGTTTTCGGAGCAACCACTGGTTTAAAGGAAAAAGT 63
DB      13 GCGGAAACACTTACTCGGAAGAGGTACTAAAGCTACCACTGGCTTCAAGATAATGCT 72
QY      64 AACGAGGAAGTGTCTCATTTGTTTTTAAATAATGAAATGTTCAACTGAAATGGAAGAAT 123
DB      73 CAAAAGAAGCATCTCTATGATTTTAAACACAAAAAGTCAATCTAAATGGAAGAAGAA 132
QY      124 ATCGATGGAATAGTTACAAAAAAGAGCTGCAGGAGCAGAGCTGAAATCTTTACACGA 183
DB      133 ATTGCTTGAATAACTACAAACAAAGATACACAGATCGGAAATGATAAAACCTACAAAGA 192
QY      184 GGAGCGAARACTTGTGGACCAAGCAGGAGCATGGAATGGGAAACCAACAGTGGATGAA 243
DB      193 GGAGCAGAAACATCATGGGACCAAGGCAACAGACATGGAATGGGAATCGAAGTCGACAGC 252
QY      244 ATGACCAAAAAGCAAGTATTCATTTTGATTTCTTTGTTTAAATAATGTTTATTTGAAGTG 303
DB      253 CTCACAAAAGCGCAAGTACTGATTTTGTCTCTCTTTTAAATAATGTTCTCTTTGAAGGT 312
QY      304 CTTAAACAAGAATAATATTTCTGGTGATGTTAATTTGTTTGTGCAACATGAATGGGA 363
DB      313 ATATTGCAAAAGAACCTTAAGTCCAAGTGACTGCTACTGTTCTTACAGCATGAACATGGT 372
QY      364 AAGACCAAGCTGGCACTGCCATGTACTAATTTGGAGGAAGAGCTTTAGTCAAGTCAA 423
DB      373 CAAGATACTGGCTATCATCTGCCATGTACTAGTGGTGAAGAAAGCTTTACAAAGCAATG 432
QY      424 GGGAAATGGTGGAGGAAGCACTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCCTGT 483
DB      433 GGAATGGTTACGAACAACTTAAACAAATTTATGGAGTAGATGGTTGAATATCAATGC 492
QY      484 AATGTGCAACTAACACCAAGCTGAAAGAAATTTAAACTAAGAGAAATAGCAGAAGAACAATGAG 543

```

```

Db      493 AAAGTACCTCTAACACCAGTTGAAAGAAATAAAATTAAGGGAATTTAGCAGAGGATGGTGAG 552
QY      544 TGGGTTACTCTACTTACTTATTAAGCATAAGCAAAACCAAAAAAGACTATATACCAAGTGTGTT 603
DB      553 TGGGATATCGTACTTAACCTACACTCACAAACAACTAAAAAACAATATACAAAATGACT 612
QY      604 CTTTTTGGAAAAATGATGTTGCTTACTATTTTTTAACTAAAAAAGAAATAAGCCTAGTCCCA 663
DB      613 CATTTTGGAAATATGATGTTGCTTACTACTCTCTAAATAAAAAAAGAAAGACAACCT----- 666
QY      664 CCAAGAGACGGAGGCTATTCTTTAGCAGTGACTCTGGCTGGGAAACTTACTTTTAAAAA 723
DB      667 GAAAGAGAGATGATGATATTATCTCAGCTCAGATTCTGGCTTCATGACAAATTTCTTAAAA 726
QY      724 GAAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGGCGGCAGAAACCGTT 783
DB      727 GAAGGCGAGAGACACTTAGTCAGTCACTTACTGAGCAATTAACCTGAAACCTGTG 786
QY      784 GAAACCAAGTAAACCACTGCGCAGGAAACTTAAGCGCGGAGAAATTTCAAACTAAAAAAGAA 843
DB      787 GAAACCAACGGTTACTACAGCTCAGGAAGTTCCCGAGGCGAGAAATACAAAACAAAAAAGAA 846
QY      844 GTTGCTATTAAAACTACACTTAAAGAGCTGTCATTAAGAGATAACCTCACCAGAGGAC 903
DB      847 GTAAGCATAAATGCAATAAGAGACTTGGTTTAATAAAAGATGTACTAGCATAGAAGGC 906
QY      904 TGGATGATGTCAGCAGACAGTACATTTGAATGATGCTCAACCCAGCTGGAGAAAAAC 963
DB      907 TGGATGATGACAGATCCAGACAGTTATATAGAAATGATGGCTCAACCCGAGGAGAAAT 966
QY      964 CTGCTGAAAAAATAGCTAGAGATTTGTACAACTCTAGCAGAAACCAAAACAGCATTT 1023
DB      967 TTAATCAAAAATACACTAGAAATAACAACCTTCTACTCTAGCAAGAAACAAAACAGCATAT 1026
QY      1024 GACTTAATTTTAGAAAAAGCTGAAACCAAGCAAACTAACCAACTTTTCACTGCCTGCACACA 1083
DB      1027 GACTTAATTTTGAAGGCAAAACCAAGCATGCTACCAACATTTAATATTAGCAATACA 1086
QY      1084 AGAACCTGCAGAAATTTTGTCTTTCATGCTGCAATCTATGTTAAAGTTTCCCATGCTATT 1143
DB      1087 AGAACATGTAATATTCAGATGCACTTGGAACTACATTTAAGTCTGCCATGCTATA 1146
QY      1144 TGTGTGT-TTTAAACAGACAGGAGGCAAAAGAAATA-----CTGTTTATTTTCATGGA 1197
DB      1147 ACTGTGTACTGAAACAGACAGGAGGAAAAAGAAATACAAATCTATTTCTATGTCATGGG 1206
QY      1198 CGAGCAGCAGCAGGCAATCTATTATTGACAGCCATAGCACAAGCAGTTGGCNAATGTT 1257
DB      1207 CCAGCATCAACAGGAAAAAGTAAATTTGCTCAAACATTTGCAAACTTAGTTGTTAATGTT 1266
QY      1258 GGTGCTATAATGCGCCCAATGTAAACTTTTCCATTTAATGACTGTACCAACAAGAACTTG 1317
DB      1267 GGTGCTACATGCGCCCAATGTAACTTTCCATTTAATGACTGTACCAANAATAAACTTA 1326
QY      1318 ATTTGGGTAGAAGAAGCTGTAACTTTTGGACAGCAAGTAAACCAAGTTTAAAGCCATTTGC 1377
DB      1327 ATATGATTTGAAGAAGCAGGAAACTTCTCTAACCAAGTAAACCAATTTCAAAGCCATATGT 1386
QY      1378 TCTGGTCAAACTATTTCGCAATGATCAAAAAGGAAAAAGGCAACAGATTTGAACCAACA 1437
DB      1387 TCAGGTCAACCAATTTAGAAATTCACAAAAGGTAAGGAGCAACAAATTTGAACCAACT 1446
QY      1438 CCACTCATCATGACCAACCAATGAGAACATTTACAGTGGTTCAGATAGGCTCGGAAGAAGA 1497
DB      1447 CTTGTAATTAATGACTACAAATGAAGACATTAATTAAGTTAGATAGGATCGGAGGAAGA 1506
QY      1498 CCAGAACACACTCAACCAATTCAGAGACAGAACTCTTAACATTTCATCTAACACATACCTTG 1557
DB      1507 CCAGAACATACACAAACCAATTAAGAGACAGAAATGTTAAACATACACCTTAAACAGAAACTG 1566
QY      1558 CTTGGTGACTTTTGGTTGTTGACAAAATGAATGGCCCATGATTTGTGCTTGGTGGTA 1617

```

1567 CCAGTGCAITTTGGCACTTTTAGAAGAACTGAATGGCCACTAATATGTGCTTGGTTGGTA 1626
1618 AGAATGGTTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCTCTGAT 1677
1627 AGAAGAGTTTACCAAGCAAAATGGCTAGTATATGTCATCAITTTGGGAAATGTACTGAT 1686
1678 TGCTCAGAAAACCTGGCGGAGCCAAAGGTGCGCAACTCTTATATAATTTACTAGGTTTCGGCA 1737
1687 TGCTCAGAAAATTTGGAGGAGCCAAAATGTCATTCCCAATAATACACCA--ACAGAC 1743
1738 CGCTCACCATTACGACACCGAAAAGTACGCTCTCAGCCAGACAACTATGCACTAATCCCA 1797
1744 TCTCAGATTTCCACATCAGTGAAGAACTTCGCCAGCGGACATCACTACGACGCAACTCCA 1803
1798 CTTCGATCGATCTCGA---GGACCTGGCTTTTAGAGCTTTGGAGCACACCAATATCTCC 1853
1804 ATACAGGAGGACCTGGATTTAGCTTTTGGAGCCGTTGGAGCGGAGCCACACACC 1862

RESULT 13

ABQ95626

ID ABQ95626 standard; DNA; 374 BP.

XX AC ABQ95626;

XX DT 28-OCT-2002 (first entry)

XX DE Tumour suppression-related oligonucleotide #1277.

XX KW Tumour; cytostatic; antiviral; neuroprotective; nootropic; neuroleptic;
XX KW tumour suppression; tumour reversion; apoptosis; viral resistance; human;
XX KW viral infection; cell degeneration disease; neurodegeneration; ds;
XX KW Alzheimer's disease; schizophrenia; immune disease; inflammatory disease.

XX OS Homo sapiens.

XX PN FR2819824-A1.

XX PD 26-JUL-2002.

XX PF 23-JAN-2001; 2001FR-00000899.

XX PR 23-JAN-2001; 2001FR-00000899.

XX PA (MOLE-) MOLECULAR ENGINES LAB SA.

XX PI Telerman A, Anson R, Tuijnder M, Susini L;

XX PS WPI; 2002-610803/66.

XX PT New nucleic acid implicated e.g. in tumor suppression, useful for
XX PT diagnosis of tumors, viral infection and cellular degeneration and for
XX PT drug screening.

XX PS Claim 1; Page 360; 623pp; French.

XX CC The present invention relates to novel human nucleic acid sequences (I).
XX CC The present sequence is one such nucleic acid sequence. Expression of (I)
XX CC are implicated in tumour suppression or reversion and apoptosis and viral
XX CC resistance. (I) are useful as probes or primers for detecting,
XX CC identifying, measuring and/or amplifying nucleic acid sequences, as
XX CC antisense reagents and for recombinant production of polypeptides. (I),
XX CC polypeptides (II) encoded by (I), vector containing (I), cells containing
XX CC these vectors and antibodies (Ab) against (II) are all useful for
XX CC treatment/prevention of viral, tumour and cell degeneration diseases
XX CC (especially neurodegeneration, such as Alzheimer's disease and
XX CC schizophrenia). Analysing the expression of (I) is also useful for
XX CC diagnosis and/or prognosis of such diseases. Transgenic animals carrying
XX CC (I) are used for studying the aetiology of these diseases (also immune
XX CC and inflammatory diseases). Note: In the present specification, SEQ ID 1
XX CC to 2280 are claimed in Claim 1, however only SEQ ID 1 to 2270 are shown
XX CC in the specification

SQ Sequence 374 BP; 116 A; 57 C; 108 G; 84 T; 0 U; 9 Other;

Query Match 15.2%; Score 307.4; DB 6; Length 374;

Best Local Similarity 88.2%; Pred. No. 6.2e-72;

Matches 329; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 174 TTACAACGAGGAGCGGAAAACCTATTGGGACCAAAAGCGAGGACATGGAAATGGGAACAC 233

DB 1 TTACAACGAGGCGGAGACCCTTTGGACCCAAAGCGAGGACATGGAAATGGGAGCGC 60

QY 234 AGTGGATGAATGACCAAAAGCAAGTATTCATTTTGGTTTAAATAATGTTT 293

DB 61 AGTGGATGACATGACCAAAAGCAAGTATTTTATTTTGGTTTAAAGTGT 120

QY 294 ATTTGAAGTGTCTTAACACAAAGATATATTTCTCGTGATGTTAATGTTTGTGCAACA 353

DB 121 GTTTGAAGTGTCTCAGACACAAAGACATAGCTCTTAGTAAATGTTACTTGGTTCTGCAGCA 180

QY 354 TGAATGGGAAAAGACCAAGGCTGCGACCTGCTACTACTAATTTGGAGGAAAGCACTTTAG 413

DB 181 TGAATGGGAAAAGACCAAGGCTGCGACCTGCTACTACTAATTTGGAGGCAAGCACTTTAG 240

QY 414 TCAAGCTCAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGTTGTT 473

DB 241 TCAACTCAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTGGAGTANATGTTG 300

QY 474 AACAGCTGTAAATGTGCAACTAACACCAAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGA 533

DB 301 GACTGCTGNAATGTNCAACTAACACCAAGCTGAAAGAAATTAACCTGNNAGAAATAGCAGA 360

QY 534 AGACATGAGTGG 546

DB 361 GGACAGTGANNGG 373

RESULT 14

ABQ94779

ID ABQ94779 standard; DNA; 421 BP.

XX AC ABQ94779;

XX DT 28-OCT-2002 (first entry)

XX DE Tumour suppression-related oligonucleotide #430.

XX KW Tumour; cytostatic; antiviral; neuroprotective; nootropic; neuroleptic;
XX KW tumour suppression; tumour reversion; apoptosis; viral resistance; human;
XX KW viral infection; cell degeneration disease; neurodegeneration; ds;
XX KW Alzheimer's disease; schizophrenia; immune disease; inflammatory disease.

XX OS Homo sapiens.

XX PN FR2819824-A1.

XX PD 26-JUL-2002.

XX PF 23-JAN-2001; 2001FR-00000899.

XX PR 23-JAN-2001; 2001FR-00000899.

XX PA (MOLE-) MOLECULAR ENGINES LAB SA.

XX PI Telerman A, Anson R, Tuijnder M, Susini L;

XX PS WPI; 2002-610803/66.

XX PT New nucleic acid implicated e.g. in tumor suppression, useful for
XX PT diagnosis of tumors, viral infection and cellular degeneration and for
XX PT drug screening.

XX PS Claim 1; Page 141; 623pp; French.

XX CC The present invention relates to novel human nucleic acid sequences (I).

CC The present sequence is one such nucleic acid sequence. Expression of (I)
CC are implicated in tumour suppression or reversion and apoptosis and viral
CC resistance. (I) are useful as probes or primers for detecting,
CC identifying, measuring and/or amplifying nucleic acid sequences, as
CC antisense reagents and for recombinant production of polypeptides. (I),
CC polypeptides (II) encoded by (I), vector containing (I), cells containing
CC these vectors and antibodies (Ab) against (II) are all useful for
CC treatment/prevention of viral, tumour and cell degeneration diseases
CC (especially neurodegeneration, such as Alzheimer's disease and
CC schizophrenia). Analysing the expression of (I) is also useful for
CC diagnosis and/or prognosis of such diseases. Transgenic animals carrying
CC (I) are used for studying the aetiology of these diseases (also immune
CC and inflammatory diseases). Note: In the present specification, SEQ ID 1
CC to 2280 are claimed in Claim 1, however only SEQ ID 1 to 2270 are shown
CC in the specification
XX
SQ Sequence 421 BP; 125 A; 77 C; 118 G; 96 T; 0 U; 5 Other;

Query Match 14.2%; Score 285.8; DB 6; Length 421;
Best Local Similarity 87.6%; Pred. No. 4.1e-66;
Matches 311; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 173 CTTTACACGAGGAGCGAACTACTTGGGACCAAGCGAGCATGGAATGGGAAACCA 232
DB 59 CTTTACACGAGGCGCGAGACCACTTGGGACCAAGCGAGCATGGAATGGGAGCG 118
QY 233 CAGTGGATGAATGACCAAAAGCAAGTATTCTTTTGGTAAATGTTT 292
DB 119 CAGTGGATGACATGACCAAAAGCAAGTATTCTTTTGGTAAAGAGTGT 178
QY 293 TATTTGAAGTCTTAACACAAAGATATATCTCTGATGATTAATGTTTGTCAAC 352
DB 179 TGTTTGAAGTGCTCAGCACAAAGAACATAGCTCCTAGTATGTTTGTCTGTCGAGC 238
QY 353 ATGAATGGGAAAGACCAAGCTGGCACTGCCATGTACTTAATTTGGAGGAAGGACTTTA 412
DB 239 ATGAATGGGAAAGACCAAGCTGGCACTGTCTATGTCATGTCATTTGGAGGACAGACTTTA 298
QY 413 GTCAAGCTCAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGTTGG 472
DB 299 GTCAACCTCAAGGAAATGTTGGAGAGGAGCTAAATGTTTACTTGGAGTATGTTGG 358
QY 473 TAACAGCTCTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAAT 527
DB 359 TGACTGCTCTAATGTGTTCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAAT 413

RESULT 15

ABQ94724

ID ABQ94724 standard; DNA; 423 BP.

AC

ABQ94724;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PI

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Telerman A, Amson R, Tuijnder M, Susini L;

WPI; 2002-610803/66.

New nucleic acid implicated e.g. in tumor suppression, useful for

diagnosis of tumors, viral infection and cellular degeneration and for

drug screening.

Claim 1; Page 125-126; 623pp; French.

The present invention relates to novel human nucleic acid sequences (I).

The present sequence is one such nucleic acid sequence. Expression of (I)

are implicated in tumour suppression or reversion and apoptosis and viral

resistance. (I) are useful as probes or primers for detecting,

identifying, measuring and/or amplifying nucleic acid sequences, as

antisense reagents and for recombinant production of polypeptides. (I),

polypeptides (II) encoded by (I), vector containing (I), cells containing

these vectors and antibodies (Ab) against (II) are all useful for

treatment/prevention of viral, tumour and cell degeneration diseases

(especially neurodegeneration, such as Alzheimer's disease and

schizophrenia). Analysing the expression of (I) is also useful for

diagnosis and/or prognosis of such diseases. Transgenic animals carrying

(I) are used for studying the aetiology of these diseases (also immune

and inflammatory diseases). Note: In the present specification, SEQ ID 1

to 2280 are claimed in Claim 1, however only SEQ ID 1 to 2270 are shown

in the specification
XX
SQ Sequence 423 BP; 125 A; 74 C; 124 G; 99 T; 0 U; 1 Other;

Query Match 14.1%; Score 284.8; DB 6; Length 423;

Best Local Similarity 89.2%; Pred. No. 7.6e-66;

Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGAGCGAACTACTTGGGACCAAGCGAGCATGGAATGGGAAACCA 232

DB 71 CTTTACACGAGGCGCGAGACCACTTGGGACCAAGCGAGCATGGAATGGGAGCG 130

QY 233 CAGTGGATGAATGACCAAAAGCAAGTATTCTTTTGGTAAATGTTT 292

DB 131 CAGTGGATGACATGACCAAAAGCAAGTATTCTTTTGGTAAAGAGTGT 190

QY 293 TATTTGAAGTCTTAACACAAAGATATATTTCTCTGATGATTAATGTTTGTCAAC 352

DB 191 TGTTTGAAGTGCTCAGCACAAAGACATAGCTCCTAGTATGTTTGTCTGTCAGC 250

QY 353 ATGAATGGGAAAGACCAAGCGCTGGCACTGCCATGTACTTAATTTGGAGGAAGGACTTTA 412

DB 251 ATGAATGGGAAAGACCAAGCGCTGGCACTGTCTATGCTGATTTGGAGGACAGACTTTA 310

QY 413 GTCAAGCTCAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGTTGG 472

DB 311 GTCAACCTCAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTTGGAGTATGTTGG 370

QY 473 TAACAGCTCTAATGTTGCAACTAACACAGCTGAAAGAAATTA 516

DB 371 TGACTGCTCTAATGTTTCAACTAACACAGCTGAAAGAAATTA 414

Search completed: January 22, 2005, 11:49:41

Job time : 958.5 secs

Tumour suppression-related oligonucleotide #375.

Tumour; cytostatic; antiviral; neuroprotective; nootropic; neuroleptic;
tumour suppression; tumour reversion; apoptosis; viral resistance; human;
viral infection; cell degeneration disease; neurodegeneration; ds;
Alzheimer's disease; schizophrenia; immune disease; inflammatory disease.

Homo sapiens.

FR2819824-A1.

26-JUL-2002.

23-JAN-2001; 2001FR-00000899.

23-JAN-2001; 2001FR-00000899.

(MOLE-) MOLECULAR ENGINES LAB SA.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 19, 2005, 16:54:30 ; Search time 164.5 Seconds
(without alignments)
8723.907 Million cell updates/sec

Title: US-10-069-056-4
Perfect score: 2019
Sequence: 1 atggctggaactgttactc.....agccgtgaacttgactaa 2019

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 35539441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1040.4	51.5	5049	1	US-08-336-345-1
2	1040.4	51.5	5049	1	US-08-336-345-2
3	1040.4	51.5	5049	2	US-08-647-655-1
4	1040.4	51.5	5049	2	US-08-647-655-2
5	109.4	5.4	4680	1	US-08-254-358-1
6	109.4	5.4	4680	1	US-08-475-391-1
7	109.4	5.4	4680	2	US-08-709-609-1
8	109.4	5.4	4680	5	PCT-US95-07178-1
9	109.4	5.4	4681	4	US-09-807-802A-18
10	109.4	5.4	4910	2	US-08-331-384-2
11	109.4	5.4	4910	3	US-08-836-087-2
12	109.4	5.4	4910	3	US-09-246-320-2
13	109.4	5.4	4910	3	US-09-546-728-2
14	109.4	5.4	7214	4	US-09-438-268-1
15	109.4	5.4	7557	4	US-09-770-315-3
16	109.4	5.4	8151	4	US-09-438-268-2
17	109.4	5.4	8179	4	US-09-438-268-5
18	109.4	5.4	8698	4	US-09-770-315-2
19	104.6	5.2	939	4	US-09-532-594B-12
20	104.6	5.2	1197	4	US-09-532-594B-13
21	104.6	5.2	1611	4	US-09-532-594B-14
22	104.6	5.2	1872	4	US-09-532-594B-3
23	104.6	5.2	1872	4	US-09-532-594B-15
24	104.6	5.2	4767	4	US-09-532-594B-1
25	99.8	4.9	969	4	US-09-807-802A-10
26	99.8	4.9	1200	4	US-09-807-802A-8
27	99.8	4.9	1641	4	US-09-807-802A-6

28	99.8	4.9	1872	4	US-09-807-802A-4	Sequence 4, Appli
29	99.8	4.9	4683	4	US-09-807-802A-19	Sequence 19, Appli
30	99.8	4.9	4718	4	US-09-807-802A-1	Sequence 1, Appli
C 31	64.8	3.2	7218	1	US-08-232-463-14	Sequence 14, Appli
C 32	61.2	3.0	1141	4	US-09-806-708B-22	Sequence 22, Appli
33	44	2.2	1141	4	US-09-806-708B-22	Sequence 22, Appli
34	40.2	2.0	277	3	US-09-007-005-3	Sequence 3, Appli
35	40.2	2.0	277	3	US-09-244-796-3	Sequence 3, Appli
C 36	39.2	1.9	1664976	4	US-08-916-421B-1	Sequence 1, Appli
C 37	39.2	1.9	1664976	4	US-09-692-570-1	Sequence 1, Appli
38	38.6	1.9	116592	4	US-09-818-512-3	Sequence 14265, A
39	38.4	1.9	1367	4	US-09-270-767-14265	Sequence 1, Appli
40	38.4	1.9	6328	3	US-08-913-832A-1	Sequence 1, Appli
41	38.4	1.9	6328	4	US-09-249-181A-1	Sequence 1, Appli
42	38.4	1.9	6328	4	US-09-158-707-1	Sequence 1, Appli
43	38.4	1.9	6475	4	US-09-620-312D-325	Sequence 325, App
44	37.8	1.9	399	4	US-09-621-976-8976	Sequence 8976, Ap
45	37.8	1.9	832	4	US-09-621-976-2813	Sequence-2813, Ap

ALIGNMENTS

RESULT 1
US-08-336-345-1
; Sequence 1, Application US/08336345
; Patent No. 5814510
; GENERAL INFORMATION:
; APPLICANT: Parrish, Colin R.
; APPLICANT: Gruenberg, Allen
; APPLICANT: Carmichael, Leland E.
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,345
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Gordon, Jennifer
; REGISTRATION NUMBER: 30753
; REFERENCE/DOCKET NUMBER: 7937-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-3090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Parvovirus
; US-08-336-345-1

Query Match 51.5%; Score 1040.4; DB 1; Length 5049;
Best Local Similarity 71.4%; Pred. No. 5.5e-291;
Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;
QY 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTTCGGAGCAACCAACTGTTAAAGGAAAA 60

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Pennie & Edmonds
;; STREET: 1155 Avenue of the Americas
;; CITY: New York
;; STATE: NY
;; ZIP: 10036-2711
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/336,345
;; FILING DATE:
;; CLASSIFICATION: 424
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Gordon, Jennifer
;; REGISTRATION NUMBER: 30753
;; REFERENCE/DOCKET NUMBER: 7937-006
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 790-9090
;; TELEFAX: (212) 869-8864/9741
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 5049 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: both
;; TOPOLOGY: unknown
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; ORIGINAL SOURCE:
;; ORGANISM: Parvovirus
;; US-08-336-345-2

Query Match 51.5%; Score 1040.4; DB 1; Length 5049;

Best Local Similarity 71.4%; Pred. No. 5.5e-291;
Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;

QY	1	ATGCTCGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTTAAAGGAAAA	60
DB	273	ATGCTCGGCAACCACTATCTAGGAAAGTTATGGAGGAGTAATTTGGTTAAAGAAACAT	332
QY	61	AGTAACCAAGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAA	120
DB	333	GCAGAAAAATGAAGCAATTTCTGTTTGTATTTAAATGTGACAACTGCTCAACTAAATGGAAG	392
QY	121	GATATCGGATGGAATAGTTTACAAAAAGAGCTCGAGGAGCAAGCTGAAATCTTTACAA	180
DB	393	GATGTTGCTGGAACTATATACCAAACTATACCAAACTTCAAAATGAAGAGCTAACATCTTTAAT	452
QY	181	CGAGGAGCGAAACTACTTGGGACCAAGC-----GAGGACATGGAATGGGAAACCA	234
DB	453	AGAGGAGCAAAACAGCAATGGATCAACCCGAAGAGAAATGAGCTGGGAATCGAA	512
QY	235	GTGGATGAATGATCAAAAAAGCAAGTATTCATTTTGTGTTTAAATAATGTTTA	294
DB	513	GTGATAGTCTCGCCAAAAGCAAGTACAACTTTTGTGATGATTAATTTAAATAATGCTT	572
QY	295	TTTGAAGTGTCTTAACCAAGAATATATTTCTGGTGTGATGTTAATTTGGTTTGGCAACAT	354
DB	573	TTTGAAGTCTTGTCTTAAAAATATAGAACCAAAATGAATGTGTTTGGTTTATTCACAT	632
QY	355	GAATGGGAAAGCAACCAAGCTCGCACTGCATGCTAATTTGGAGAAAGGACTTTAGT	414
DB	633	GAATGGGAAAGATCAAGCTGGCATTTGTCATGTTTCTTCATAGTAAGAACTTACAA	692
QY	415	CAAGCTCAAGGGAATGGTGGAGAGGCAACTAAATGTTTACTCGAGCAGATGGTTGGTA	474
DB	693	CAAGCAACTGTAATGGCTTACGACAGCAATGAATATGATTTATGGAGTAGATGTTGGTG	752
QY	475	ACAGCCTGTAATGTGCACTTAAACCAAGCTGAAAGAAATTTAACTAGAGAAATAGCAGAA	534

DB	753	ACTCTTTGTTGGTAAACTTTAACACCACTGAAAGATTAAAGCTCAGAGAAATTCAGAA	812
QY	535	GAACAATGAGTGGGTACTCTACTTACTTATAGCATAGCAAAACCAAAAGACTATACC	594
DB	813	GATAGTGAATGGGTGACTATATTAACATACAGACATAAGCAAAACAAAAAGACTATGTT	872
QY	595	AAGTGTGTTCTTTTGGAAACATGATTGCTTACTATTATTTTAACTTAAAGAAATAAGC	654
DB	873	AAAAATGGTTCTTTTGGAAATATGATAGCATATTTCTTTTAAAGAAAAAATTTGTC	932
QY	655	ACTAGTCCACCAAGAGACGCGAGCTATTTCTTAGCAGTACTCTCGCTGGAAAACTAAC	714
DB	933	CACAT---GACAAAAGAAAGTGGCTATTTTAAAGTACTGATTTCTGTTGGAAATTTAAC	989
QY	715	TTTTTAAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATCGCGCCA	774
DB	990	TTTATGAAGTATCAAGACAGACAAATTTGTGAGCACACTTTTACACTGAACAAATGAACCA	1049
QY	775	GAAACGGTTGAAACCACTGCGCGAGGAACTAAGCGCGCAGAAATTTCAAACT	834
DB	1050	GAAACGGTTGAAACCACTGCGCGAGGAACTAAGCGCGCAGGAAATTTCAAACT	1109
QY	835	AAAAAGAAAGTTGCTTATTTAAACTACACTTTAAAGAGCTGCTGATATAAGAGTAACTCA	894
DB	1110	AAAAAGAAAGTGTCAATCAATGCTACTTTGCGGACTTGGTTAGTAAAGAGTAACTCA	1169
QY	895	CCAGGAGTGGATGATGATGAGCAGCAGACAGTTACATTGAAATGATGCTCAACAGGT	954
DB	1170	CCTGAAGACTGGATGATGTTTACAAACAGATAGTTATTTGAAATGATGCGCAACAGGA	1229
QY	955	GGAGAAACCTGCTGAAAAATACGCTAGAGATTTGTACATACTCTAGCCAGAACCAA	1014
DB	1230	GGTGAATCTTTTAAAAAATACACTTGAATTTGTACTTTGACTTTAGCAAGAACAAA	1289
QY	1015	ACAGCAATTTGACTTAAATTTAGAAAAAGCTGAAACCAAGCACTAACCACTTTTCACTG	1074
DB	1290	ACAGCAATTTGAAATTAATCTTGAAGAAAGAGATATCTAACTACTACTTTGATCTT	1349
QY	1075	CCTGACACAAAGAACTGCAAGATTTTGTCTTTTATGGCTGGAACTATGTTAAAGTTGC	1134
DB	1350	GCAATTTCTAGAACATGTCAATTTTGTAGATGCAAGGATGGAATTTGATTTAAAGTTGT	1409
QY	1135	CATGCTATTTGCTGTGTTTAAACAGACAAAGGAGCAAAAGAAATCTGTTTATTTTCAT	1194
DB	1410	CACGCTATAGCATGTTTAAATAGCAAGGTTGTAAGAAATACAGTCTCTTTTTCAT	1469
QY	1195	GGACCAAGCAGACAGCAAGCAATCTATTATTGCAAGCCATAGCAAGCAGTGTGGCAAT	1254
DB	1470	GGACCAAGCAGTACAGGAAATCTATCATTTGCTCAAGCCATAGCAAGCTGTGGTAAAT	1529
QY	1255	GTGTTGCTGATTAATGCAAGCAATGTAAACTTTTCCATTTTAACTGTACCAACAGAAC	1314
DB	1530	GTGTTGTTTATATGCAAGCAATGTAAATTTTCCATTTTAACTGTACCAATAAAT	1589
QY	1315	TTGATTTGGTGTAGAAAGCTGTAACTTTTGGACAGCAAGTAAACAGTTTAAAGCCAT	1374
DB	1590	TTAAATTTGGATTTGAAGAGCTGGTAACTTTTGGTCAACAAAGTTAATCAATTTAAAGCAATC	1649
QY	1375	TGCTCTGCTCAACTATTGCAATTTGATCAAAAGGAAAGGCAAGCAAGATTTGAACCA	1434
DB	1650	TGTTCTGGACAAACCAATTTAGATTTGATCAAAAGGTAAGAAAGTAAGCAATTTGAACCA	1709
QY	1435	ACACCACTCATCATGACCAAAATGAGAACTATTAAGTGGTGTGAGATAGCTGCGAGAA	1494
DB	1710	ACTCCAGTAAATTTAGCAACTAATGAATATAAATAATTTGTGAGATTTGGATGTGAAGAA	1769
QY	1495	AGACCAAGACACACTCAACCAATCAGACAGAGATGCTTAACTTCTTCACTTAAACATACC	1554
DB	1770	AGACCTGAACATACACAACTAATAGACAGCAAGTGTGAAACATTAAGTTAGTATGAAG	1829
QY	1555	TTGCTGCTGACTTGTGTTTGGTTGTAAGAAATGAATGCCCCATGATTTGCTGCTGTTG	1614
DB	1830	CTTCCAGGAGACTTTGGTTGGTTTGAAGAAAGTGGCTTTTAAATATATGTCATGGTTA	1889

1615 GTAAAGATGGTTACCAATCTTACCATGGCAAGCTACTGTCTAAATGGGCAAGTTCTCT 1674
1890 GTTAAACATGGTTTGAATCAACCATGGCTAACTATACACATCATTTGGGGAAGTACCA 1949
1675 GATTGGTCAGAAACTGGGGGAGCCAAAGGTGCCAACTCTTATAAATTTACTAGTTGCG 1734
1950 GAATGGGATGAAACTGGGGGAGCCCTAAAATACAGAAGGTATATAATTCACCGGTTGC 2009
1735 GCAGCTCACCATTTCAGCACACGAAAGTAAGCTCTCAGCGAGACTATGCACTAACT 1794
2010 AAAGACT---TAGAGACACAGCGGCAAGCAATCTCAGAGTCAAGACCAAGTTCTAACT 2066
1795 CCACTTCGATCGGATCTCGAGGAGCTGGCTTTAGAGCCTTTGGAGCACACCAAAATATCTCT 1854
2067 CCTCTGACTCGGAGCTAGTGGACCTTGCATGGAACCGTGGAGTACTCCAGATACGCTT 2126
1855 GTTGGGGGACTGTCAGAAACCCAGAACACTCTGGGGAAGCTGGTTCGAAAGCTCCCAAGAT 1914
2127 ATTGCAGAAACTGCAATCAACATCAACCAACTTGGCGTTACTCA-----CAAAGAC 2180
1915 GGTCAACTGAGCCCAACTCTGTCAGAGATCGAGAGGATTTGAGAGGTGCTTTCGGTGG 1974
2181 GTGCAACGAGTCCGAGTGGTTCGAATAGAGCGACCTGAGAGCCATCTTTACTTCT 2240
1975 GAACCGTTGAAGAAGACTTTCAGCGA 2000
2241 GAACAATTGGAAGAGATTTTCGAGA 2266

RESULT 3
US-08-647-655-1
; Sequence 1, Application US/08647655
; Patent No. 5885585
; GENERAL INFORMATION:
; APPLICANT: Parrish, Colin R.
; APPLICANT: Gruenberg, Allen
; APPLICANT: Carmichael, Leland E.
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/647,655
; FILING DATE: On Even Date Herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gordon, Jennifer
; REGISTRATION NUMBER: 30,753
; REFERENCE/DOCKET NUMBER: 7937-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORGANISM: Parvovirus

US-08-647-655-1

Query Match 51.5%; Score 1040.4; DB 2; Length 5049;
Best Local Similarity 71.4%; Pred. No. 5.5e-291;
Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;
QY 1 ATGGCTGGAAATGCTTACTCTGATGAAGTCTTGGGAGCAACCAACTGGTTAAAGGAAAAA 60
Db 273 ATGTCTGGCAACCCAGTATATCTGAGGAAGTTATGGAGGGAGTAAATTTGGTTAAAGAAACAT 332
QY 61 AGTAACCCAGGAAGTGTCTCATTTGTTTAAAAATGAAAATGTTCAACTGAATGAAAAA 120
Db 333 GCAGAAATGAAAGCAATTTTCGTTGTTTAAATGTGACACGTCCTCACTAAATGGAAG 392
QY 121 GATATCGGATGAATAGTTTACAAAAAGAGCTGCAAGGAGCAGAGCTGAAATCTTTTACAA 180
Db 393 GATGTTGCTGGTGAACCACTATACCAAA CCAATTCAAAATGAAGAGCTACATCTTTTAAT 452
QY 181 CGAGGAGCGGAAACTACTTTGGGACCAAGC-----GAGGACATGGAATGGGAAACCA 234
Db 453 AGAGGAGCACAAACAGCAATGGATCAACCCGAAAGAAAGAAATGGACTGGGAATCGGAA 512
QY 235 GTGGATGAATGACCAAAAAGCAAGTATTCATTTTGAATCTTTGGTTAAAAAATGTTTA 294
Db 513 GTTGATAGTCTCTCGCAAAAAGCAAGTACAAACTTTTGATGCAATTAATTTAAAAAATGTCT 572
QY 295 TTTGAAGTGTCTTAAACACAAAGAAATATATTTCTGTTGATGTTAAATTTGGTTTCTGCAACAT 354
Db 573 TTTGAAGTCTTTGTTCTTAAAAATATAGAACCAATGAATGTGTTGTTTATTCACAT 632
QY 355 GAATGGGAAAAAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAGACTTTTACT 414
Db 633 GAATGGGAAAAAGATCAAGGCTGGCATTGTCTGTTTACTTTCATAGTAAGAACTTACAA 692
QY 415 CAAGCTCAAGGAAATGTGGAGAGGCACTAAATGTTTACTTGGAGCAGATGTTGGTA 474
Db 693 CAAGCAACTGTTAAATGGCTACGACACAAATGAAATATGTATTGGAGTAGATGGTTGGTG 752
QY 475 ACAGCTGTAATGTGCACTCAACACAGCTGAAAGAAATTTAAACTAAGAGAAATAGCAGAA 534
Db 753 ACTCTTTGTTGCGTAACTTAAACCACTGAAAGATTTAGCTCAGAGAAATTCGAGAA 812
QY 535 GACAATGAGTGGTTTACTCTACTTATTATAGCATAAGCAAAACCAAAAGACTATACC 594
Db 813 GATAGTGAATGGTGTACTATATTAACATACAGACATAAGCAAAACCAAAAGACTATGTT 872
QY 595 AAGTGTGTTCTTTTGGAAACATGATGCTTACTATTTTAACTTAAAGAAATATAAGC 654
Db 873 AAAATGGTTTCATTTTGGAAATATGATAGCATATTACTTTTAAACAAAGAAAAATTTGTC 932
QY 655 ACTAGTCCACCAAGAGAGCGAGGCTATTTCTTAGCAGTACTCTGGCTGGAAAACTAAC 714
Db 933 CACAT---GACAAAAGAAAGTGGCTATTTTTTAAGTACTGATTTCTGGTTGGAATTTAAC 989
QY 715 TTTTAAAGAAAGGCGAGCGCCATCTAGTGAACAACTATACATCTGATGACATCGGGCCA 774
Db 990 TTTATGAAGTATCAAGACAGACAAATTTGTACACACACTTTACACTGAAACAAATGAAACCA 1049
QY 775 GAAACGGTTGAAACCAACAGTAAACCACTGGGAGGAACTAAGCGCGGAGAAATTCAACT 834
Db 1050 GAAACCGTTGAAACCAACAGTAAACCACTGGGAGGAACTAAGCGCGGAGAAATTCAACT 1109
QY 835 AAAAAAGAAAGTTGCTATTAAAACTACACTTAAAGAGCTGGTGCATATAAAGAGTAACTCA 894
Db 1110 AAAAGGAAGTGTCAATCAATGTTACTTTTGGGACTTGGTTAGTAAAGAGTAACTCA 1169
QY 895 CCAGAGGACTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 954
Db 1170 CCTGAAGACTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1229
QY 955 GGAGAAACCTGCTGAAAAATAGCGTAGAGATTTGCTACTACTACTACTACTACTACTACT 1014
Db 1230 GGTGAAATCTTTTAAAAAATACACTTGAATTTTGTACTTTGACTTTTGTAGCAAGAACAAA 1289

```
QY 1015 ACAGCATTTGACCTTAATTTAGAAAAGCTGAAACGACGAAACCTAACCAACTTTTCACTG 1074
Db 1290 ACAGCATTTGAAATTAATCTTGAAGAAGCAGATAATACTAACTAACTTTGATCTT 1349
QY 1075 CTTGACACAGAACTCGCAAAATTTTGGCTTTTTCATGGCTGGAACCTATGTTAAAGTTTGC 1134
Db 1350 GCAAAATCTAGAACATGTCAAATTTTGTAGATGACGGATGGATGGAATTAAGTTTGT 1409
QY 1135 CATGCTATTTGCTGTGTTTTTAAACAGACAAGGAGGCAAAAGAAATACTGTTTTTATTCAT 1194
Db 1410 CACGCTATAGCATGTGTTTTTAAATAGACAAGGTGGTAAAGAAATACAGTTCTTTTTTCAT 1469
QY 1195 GGAACGACGACAGCAAACTATTTATTTGCAACAGCCATAGCACAGCAGTTGGCAAT 1254
Db 1470 GGACCAAGCAAGTACAGAAATCTATCATTTGCTCAAGCCATAGCACAGCTGTGGTAAT 1529
QY 1255 GTTGGTGTCTATATGACGCAATGTAAACTTTTCCATTTTAACTGCTACCAACAAGAAC 1314
Db 1530 GTTGGTGTCTATATGACGCAATGTAAATTTTCCATTTTAACTGCTACCAATAAAT 1589
QY 1315 TTGATTTGGGTAGAAGCTGGTAACTTTTGGACAGCAAGTAAACCACTTTAAAGCCATT 1374
Db 1590 TTAATTTGGATTGAAGCTGGTAACTTTTGGTCAACAAGTTAATCAATTTAAAGCAATC 1649
QY 1375 TGCTCTGGTCAAACTATTCGCATTTGATCAAAAAGGAAAGGCAAGCAAAAGATTTGAACCA 1434
Db 1650 TGTCTCGGACAAACAATTAGAATTTGATCAAAAAGGTAAGGAAGTAAAGCAAAATTTGAACCA 1709
QY 1435 ACACCACTCATATGACCAACATGAACTTACAGTGTGAGATAGCTGCGGAAGAA 1494
Db 1710 ACTCCAGTAATATGACCACTAATGAAATATAACAATTTGGAATTTGATGTGAAGAA 1769
QY 1495 AGACCAAGAACACTCAACCAATCAGACAGCAAGTCTTAACTTCACTTAAACACATACC 1554
Db 1770 AGACTTGAACATACACAACTAAGACAGAGATGTTGACATTAAGTTAGTATGTAG 1829
QY 1555 TTCCCTGGTGACTTTGGTTGGTTGACAAAATGAAATGGCCCAATGATTTGCTGTGGTTG 1614
Db 1830 CTTCCAGGAGACTTTGGTTGGTTGATAAAGAAAGTGGCTTTAATATGTGCATGTTA 1889
QY 1615 GTAAAGATGGTTACCAATCTACCTGCGCAAGCTACTGTCTAAATGGGCAAGTTCTCT 1674
Db 1890 GTTAAACATGGTTTGAATCAACCATGCTGCTAACTATACATCATTTGGGAAAGTACCA 1949
QY 1675 GATTGGTCAGAAAACCTGGGCGGAGCCAAAGTGCCTTCTATAAATTTACTAGGTTGC 1734
Db 1950 GAATGGATGAAACTGGGCGGAGCTTAAATACAGAAAGTTAATTCACCAAGTTGC 2009
QY 1735 GCAGCTCACCATTTACGACACCGGAAAGTACGCTCTCAGCAGAACTATGCACTAACT 1794
Db 2010 AAAGACT---TAGAGACACAAGCGGCAAGCAATCCTCAGAGTCAAGACCAAGTTCTAACT 2066
QY 1795 CCATTTGCATCGGATCTGAGGACCTGGCTTTTAGAGCTTGGAGCACCAACCACTCTCT 1854
Db 2067 CCTCTGATCCGACGAGTGGGACCTTGCACTGGAACCGGAGTACTCCAGATACGCTT 2126
QY 1855 GTTGGCGGCACTGACAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCGCAAGAT 1914
Db 2127 ATTGCAAACTGGAATCAACATCAACCAACTTGGCGTTACTCA-----CAAGAC 2180
QY 1915 GGTCAACTGAGCCCACTTGGTCAGAGATCAGGAGGATTTGAGAGCGTGTTCGGTGGC 1974
Db 2181 GTCAAGCGAGTCCGACGTGGTCGAAATAGAGGACAGCTGAGAGCGCATCTTTACTTCT 2240
QY 1975 GAACCGTTGAAGAAAGACTTCAGGA 2000
Db 2241 GAACAATTTGAAGAAAGATTTTCGAGA 2266
```

RESULT 4

US-08-647-655-2

; Sequence 2, Application US/08647655

```
; Patent No. 5895585
; GENERAL INFORMATION:
; APPLICANT: Parrish, Colin R.
; APPLICANT: Gruenberg, Allen
; APPLICANT: Carmichael, Leland E.
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/647,655
; FILING DATE: On Even Date Herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gordon, Jennifer
; REGISTRATION NUMBER: 30,753
; REFERENCE/DOCKET NUMBER: 7937-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Parvovirus
; US-08-647-655-2
```

```
Query Match 51.5%; Score 1040.4; DB 2; Length 5049;
Best Local Similarity 71.4%; Pred. No. 5.5e-291;
Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;

QY 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTTAAAGGAAAA 60
Db 273 ATGCTGGCAACCACTATCTAGGAAGTTATGGAGGAGTAATTTGGTTAAGAAACAT 332
QY 61 AGTAACGAGGAAGTGTCTCATTTGTTTTTAAATAAGAAATGTTCAACTGAATGAAAA 120
Db 333 GCAGAAAAATGAAGCAATTTTCGTTTGTAAATGTGCAACCGTCCAACCTAAATGGAAG 392
QY 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTCAGAGGAGCAGCTGAATCTTTACAA 180
Db 393 GATGTTTCGTGGCAACACTATACCAACCAATTCAAAAATGAAGAGCTAACTCTTTAAT 452
QY 181 CGAGGAGCGGAATACTACTTTGGGACCAAGC-----GAGGACATGGAATGGGAAACCA 234
Db 453 AGAGGAGCAACACAGCAATGGATCAACCGAAGAGAAATGAGCTGGGAATCGGA 512
QY 235 GTGGATGAATGACCAAAAAAGCAAGTATTCTTTTGTGTTTGTGTTTAAAAAATGTTTA 294
Db 513 GTTGATAGTCTCGCCAAAAAGCAAGTACAACTTTTGTGATGATTAATTAATAAATGTCTT 572
QY 295 TTTGAAGTGTCTTAACCAAGAAATATTTCTCGGTGATGTTAATTTGGTTTGTGCAACAT 354
Db 573 TTTGAAGTCTTTGTTTCTAAAAATATAGAACCAAAATGAATGTGTGTTTGTGTTTATCAACAT 632
QY 355 GATGGGAAAGACCAAGCTGGCACTGCATGTACTAATTTGGAGAAAGGACTTAGT 414
Db 633 GAATGGGAAAGATCAAGCTGGCATTGTGATGTTTACTTCTCATAGTAAGAACTTACAA 692
```



```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-254-358-1

Query Match          5.4%; Score 109.4; DB 1; Length 4680;
Best Local Similarity 53.7%; Pred. No. 4e-21;
Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

Qy 1168 GGCAGGAGGAAATCTGTTTATTTCATGGACGAGCCAGCAGCAGCAATCTATTATTGCA 1227
Db      |||||
Qy 1293 GGCAGGAGGAAATCTGTTTATTTCATGGACGAGCCAGCAGCAGCAATCTATTATTGCA 1352
Db      |||||
Qy 1228 CAAGCCATAGCAGCAAGCAGTTGGCAATGTTGGTGTCTATATGCAGCAGCAATGTAACCTTT 1287
Db      |||||
Qy 1353 GAGGCCATAGCCACACTGTCGCCCTTCTACGGGTGCGTAAACTGGACCAATGAGAACTTT 1412
Db      |||||
Qy 1288 CCATTTAATCACTGTACCAACAGAACTTGAATTTGGTGTCTATATGCAGCAGCAATGTAACCTTT 1347
Db      |||||
Qy 1413 CCCTTCAACGACTGTGTCGACAGATGGTGTATCTGGTGGAGGAGGGAAGATGACCGCC 1472
Db      |||||
Qy 1348 CAGCAAGTAAACCAAGTTTAAAGCAATTTGCTCTGGTCAAACTATTTCGCAATTTGTAACAAA 1407
Db      |||||
Qy 1473 AAGTCTGTGAGTGGCCAAAGCCATTTCTCGAGGAGCAGGTGCGGTGGACCAAGAA 1532
Db      |||||
Qy 1408 GGAAGAGGAGCAAAACAGATTGAACCAACAGATCATATGACCAAAATGAGAACTTT 1467
Db      |||||
Qy 1533 TGCAAGTCTCTCGGCCAGATAGACCCGACTCCCGTGATGTCACCTCCAACACCAATG 1592
Db      |||||
Qy 1468 ACAGTGTGAGATAGCTCGGAGAAAGACCAAGACCACTCAACCAATCAGAGACAGA 1527
Db      |||||
Qy 1593 TGGCCGTGTGATTGACGGGAATCAACGACCTTCGACACACAGCAGCGGTTGCAAGACCGG 1652
Db      |||||
Qy 1528 ATGCTTAAACATTCATCTAAACATACCTTCGCTGGTGACTTTGTTTGTGACAAAT 1587
Db      |||||
Qy 1653 ATGTTCAAATTTGAATCAACCCCGCTCTGATCATGACTTTGGGAGGTCACCAAGCAG 1712
Db      |||||
Qy 1588 GAA 1590
Db      |||
Qy 1713 GAA 1715
Db      |||
```

RESULT 6

```
US-08-475-391-1
; Sequence 1, Application US/08475391
; Patent No. 5786211
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,391
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/254,358
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5786211and, Greta E.
```

```
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31975
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-475-391-1
```

```
Query Match          5.4%; Score 109.4; DB 1; Length 4680;
Best Local Similarity 53.7%; Pred. No. 4e-21;
Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

Qy 1168 GGCAGGAGGAAATCTGTTTATTTCATGGACGAGCCAGCAGCAGCAATCTATTATTGCA 1227
Db      |||||
Qy 1293 GGCAGGAGGAAATCTGTTTATTTCATGGACGAGCCAGCAGCAGCAATCTATTATTGCA 1352
Db      |||||
Qy 1228 CAAGCCATAGCAGCAAGCAGTTGGCAATGTTGGTGTCTATATGCAGCAGCAATGTAACCTTT 1287
Db      |||||
Qy 1353 GAGGCCATAGCCACACTGTCGCCCTTCTACGGGTGCGTAAACTGGACCAATGAGAACTTT 1412
Db      |||||
Qy 1288 CCATTTAATCACTGTACCAACAGAACTTGAATTTGGTGTCTATATGCAGCAGCAATGTAACCTTT 1347
Db      |||||
Qy 1413 CCCTTCAACGACTGTGTCGACAGATGGTGTATCTGGTGGAGGAGGGAAGATGACCGCC 1472
Db      |||||
Qy 1348 CAGCAAGTAAACCAAGTTTAAAGCAATTTGCTCTGGTCAAACTATTTCGCAATTTGTAACAAA 1407
Db      |||||
Qy 1473 AAGTCTGTGAGTGGCCAAAGCCATTTCTCGAGGAGCAAGGTGCGGTGGACCAAGAA 1532
Db      |||||
Qy 1408 GGAAGAGGAGCAAAACAGATTGAACCAACAGATCATATGACCAAAATGAGAACTTT 1467
Db      |||||
Qy 1533 TGCAAGTCTCTCGGCCAGATAGACCCGACTCCCGTGATGTCACCTCCAACACCAATG 1592
Db      |||||
Qy 1468 ACAGTGTGAGATAGCTCGGAGAAAGACCAAGACCACTCAACCAATCAGAGACAGA 1527
Db      |||||
Qy 1593 TGGCCGTGTGATTGACGGGAATCAACGACCTTCGACACACAGCAGCGGTTGCAAGACCGG 1652
Db      |||||
Qy 1528 ATGCTTAAACATTCATCTAAACATACCTTCGCTGGTGACTTTGTTTGTGACAAAT 1587
Db      |||||
Qy 1653 ATGTTCAAATTTGAATCAACCCCGCTCTGATCATGACTTTGGGAGGTCACCAAGCAG 1712
Db      |||||
Qy 1588 GAA 1590
Db      |||
Qy 1713 GAA 1715
Db      |||
```

RESULT 7

```
US-08-709-609-1
; Sequence 1, Application US/08709609
; Patent No. 5858775
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/709,609
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 585877sand, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31975
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-709-609-1

Query Match 5.4%; Score 109.4; DB 2; Length 4680;
Best Local Similarity 53.7%; Pred. No. 4e-21;
Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 1168 GGCAGGAGAAATACCTGTTTATTTTCATGGACCCAGCAGCAGGCAAAATCTATTATTGCA 1227
DB 1293 GGCAGGAGAAATACCTGTTTATTTTCATGGACCCAGCAGCAGGCAAAATCTATTATTGCA 1352
QY 1228 CAAGCCATAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1287
DB 1353 GAGGCCATAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1412
QY 1288 CCATTTAATGCTGCTACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1347
DB 1413 CCCTTCAACGACTGTGTCGACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1472
QY 1348 CAGCAAGTAAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1407
DB 1473 AAGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1532
QY 1408 GGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1467
DB 1532 TGCAAGTCTCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1592
QY 1468 ACAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1527
DB 1593 TGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1652
QY 1528 ATGCTTAACATTTATCTAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1587
DB 1653 ATGTTCAAAATTTGAATCACCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1712
QY 1588 GAA 1590
DB 1713 GAA 1715
```

```
RESULT 8
PCT-US95-07178-1
; Sequence 1, Application PC/TUS9507178
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and Methods
; NUMBER OF INVENTION: 3
; CORRESPONDENCE ADDRESS:
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07178
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Noland, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31975
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
PCT-US95-07178-1

Query Match 5.4%; Score 109.4; DB 5; Length 4680;
Best Local Similarity 53.7%; Pred. No. 4e-21;
Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 1168 GGCAGGAGAAATACCTGTTTATTTTCATGGACCCAGCAGCAGGCAAAATCTATTATTGCA 1227
DB 1293 GGCAGGAGAAATACCTGTTTATTTTCATGGACCCAGCAGCAGGCAAAATCTATTATTGCA 1352
QY 1228 CAAGCCATAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1287
DB 1353 GAGGCCATAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1412
QY 1288 CCATTTAATGCTGCTACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1347
DB 1413 CCCTTCAACGACTGTGTCGACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1472
QY 1348 CAGCAAGTAAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1407
DB 1473 AAGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1532
QY 1408 GGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1467
DB 1532 TGCAAGTCTCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1592
QY 1468 ACAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1527
DB 1593 TGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1652
QY 1528 ATGCTTAACATTTATCTAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1587
DB 1653 ATGTTCAAAATTTGAATCACCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1712
QY 1588 GAA 1590
DB 1713 GAA 1715
```

```
RESULT 9
US-09-807-802A-18
; Sequence 18, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences, Vectors and Host Cells Containing Same
; FILE REFERENCE: GNVNP.031USA
```


Db 1184 ATGTTCAAAATTGAACACCCCGCTCTGGATCATGACTTTTGGGAAGGTCAACCAAGCAG 1125

Qy 1588 GAA 1590
|||

Db 1124 GAA 1122

RESULT 13
US-09-546-738-2/c
; Sequence 2, Application US/09546738
; Patent No. 6387368
; GENERAL INFORMATION:
; APPLICANT: Trustees of University of Pennsylvania
; Wilson, James M.
; Kelley, William M.
; Fisher, Krishna J.
; TITLE OF INVENTION: Hybrid Adenovirus-AAV Vector and
; Methods of Use Thereof
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, PO Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/546.738
; FILING DATE: 11-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/246,320
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GNVN.007PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4910 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-546-738-2

Query Match 5.4%; Score 109.4; DB 3; Length 4910;
Best Local Similarity 53.7%; Pred. No. 4.2e-21;
Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

Qy 1168 GGCAAGAATACTGTTTATTTTCATGGACCGCAGCAGCAGCAATCTATTATTGCA 1227
|||

Db 1544 GGCAAGAGGAACCACTCTGGCTGTTTGGCGCTGCAACTACCGGGAAGACCAACATCGCG 1485
|||

Qy 1228 CAAGCCATAGCACAAAGCAGTTGGCAATGTTGGTGTATATGCAGCAATGTAAACTTT 1287
|||

Db 1484 GAGGCCATAGCCACACTGTGCCCTTCTACGGGTGCGTAACCTGGACCAATGAGAATTT 1425
|||

Qy 1288 CCATTTAATGACTGTACCAAGAAGACTTGATTGGGTAGAGAAGCTGGTAACTTTTGA 1347
|||

Db 1424 CCCTTCAACGACTGTGTGCAAGATGATGATCTGGTGGGAGGAGGGGAAGATGACCGCC 1365
|||

Qy 1348 CAGCAAGTAAACCAAGTTTAAAGCCATTTGCTCTGGTCAAACTATTTCGCTTGTATCAAAAA 1407
|||

Db 1364 AAGTCTGTGAGTCGGCCAAAGCCATTCTCGGAGGAAGAGGTGCGCGTGGACCAGAA 1305
|||

Qy 1408 GGAAGAGGCAGCAAAACAGATTGAACCAACACCAAGTATCATGACCAAAATGAGAACTT 1467
|||

Db 1304 TGCAAGTCTCTCGGCCCCAGATAGACCGACTCCCGTGATCGTCACTCAACACCAACATG 1245
|||

Qy 1468 ACAGTGTGAGATAGCTGCGGAAGAAAGACCAAGCACTCAACCAATCAGAGACAGA 1527
|||

Db 1244 TGCGCCGTGATTGACGGGAACCTCAACGACCTTCGAACACACGAGCGCTTGGCAAGACCGG 1185
|||

Qy 1528 ATGCTTAAACATTCTTAACACATACCTTCCCTGGTGACTTTTGGTTGGTTGACAAAAAT 1587
|||

Db 1184 ATGTTCAAAATTGAACACCTACCCCGCTCTGGATCATGACTTTGGGAAGGTCAACCAAGCAG 1125
|||

Qy 1588 GAA 1590
|||

Db 1124 GAA 1122

RESULT 14
US-09-438-268-1/c
; Sequence 1, Application US/09438268
; Patent No. 6491907
; GENERAL INFORMATION:
; APPLICANT: Rabinowitz, Joseph E.
; APPLICANT: Samulski, Richard J.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING
; FILE REFERENCE: 5470-186
; CURRENT APPLICATION NUMBER: US/09/438,268
; CURRENT FILING DATE: 1999-11-10
; EARLIER APPLICATION NUMBER: 60/107,840
; EARLIER FILING DATE: 1998-11-10
; EARLIER APPLICATION NUMBER: 60/123,651
; EARLIER FILING DATE: 1999-03-10
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 7214
; TYPE: DNA
; ORGANISM: Virus
US-09-438-268-1

Query Match 5.4%; Score 109.4; DB 4; Length 7214;
Best Local Similarity 53.7%; Pred. No. 5.2e-21;
Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

Qy 1168 GGCAAGAATACTGTTTATTTTCATGGACCGCAGCAGCAGCAATCTATTATTGCA 1227
|||

Db 3899 GGCAAGAGGAACCACTCTGGCTGTTTGGCGCTGCAACTACCGGGAAGACCAACATCGCG 3840
|||

Qy 1228 CAAGCCATAGCACAAAGCAGTTGGCAATGTTGGTGTATATGCAGCAATGTAAACTTT 1287
|||

Db 3839 GAGGCCATAGCCACACTGTGCCCTTCTACGGGTGCGTAACCTGGACCAATGAGAATTT 3780
|||

Qy 1288 CCATTTAATGACTGTACCAAGAAGACTTGATTGGGTAGAGAAGCTGGTAACTTTTGA 1347
|||

Db 3779 CCCTTCAACGACTGTGTGCAAGATGATGATCTGGTGGGAGGAGGGGAAGATGACCGCC 3720
|||

Qy 1348 CAGCAAGTAAACCAAGTTTAAAGCCATTTGCTCTGGTCAAACTATTTCGCTTGTATCAAAAA 1407
|||

Db 3719 AAGTCTGTGAGTCGGCCAAAGCCATTCTCGGAGGAAGCAAGGTGCGCGTGGACCAAGAA 3660
|||

Qy 1408 GGAAGAGGCAGCAAAACAGATTGAACCAACACCAAGTATCATGACCAAAATGAGAACTT 1467
|||

Db 3659 TGCAAGTCTCTCGGCCCCAGATAGACCGACTCCCGTGATCGTCACTTCAACCAACATG 3600
|||

Qy 1468 ACAGTGTGAGATAGGCTGCGAAGAAAGACCAAGCACTCAACCAATCAGAGACAGA 1527
|||

Db 3599 TGGCGCGTGTGACCGGGAACCTCAACGACCTTCGAACACACGAGCGCTTGGCAAGACCGG 3540
|||

Qy 1528 ATGCTTAAACATTCTTAACACATACCTTGGCTGGTGACTTTGGTTGGTTGACAAAAAT 1587
|||

Db 3539 ATGTTCAAAATTGAACCTACCCGCGTCTGGATCATGACTTTGGGAAGGTCAACCAAGCAG 3480
Qy 1588 GAA 1590
Db 3479 GAA 3477

RESULT 15

US-09-770-315-3
; Sequence 3, Application US/09770315
; Patent No. 6429001
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; TITLE OF INVENTION: Recombinant AAV Packaging Systems
; FILE REFERENCE: 20263-501
; CURRENT APPLICATION NUMBER: US/09/770,315
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,536
; PRIOR FILING DATE: 2000-01-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 7557
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: recombinant DNA
US-09-770-315-3

Query Match 5.4%; Score 109.4; DB 4; Length 7557;
Best Local Similarity 53.7%; Pred. No. 5.3e-21;
Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;
Qy 1168 GGCMAAAGAAATACTGTTTATTTTCATGACACGCCGACGACAGGCAAACTATATTGCA 1227
Db 1321 GGCMAAGAGGAACCACTCTGGCTGTTTGGGCTGCAACTACCGGGAAGACCAACATCGCG 1380
Qy 1228 CAAGCCATAGCACAGCAGTTGGCAATGTTGGTTGCTATATAATGCACCAATGTAACTTT 1287
Db 1381 GAGGCCATAGCCACACTGTCCTTCTACGGTGGCTAACTGGACCAATGAGAACTTT 1440
Qy 1288 CCATTTAATGACTGTACCAACAGAACTTGATTTGGGTAGAAGAAGCTGGTAACTTTGGA 1347
Db 1441 CCCTTCAACGACTGTGTCGACAAGATGGTGATCTGGTGGGAGGAGGAGATGACCGCC 1500
Qy 1348 CAGCAAGTAAACCAAGTTTAAAGCAATTTGCTCTGGTCAAACTATTGCAATTGATCAAAA 1407
Db 1501 AAGGTCGTGGAGTCGGCCAAAGCCATTCTCGGAGGAAGCAAGGTGGCGGTGGACCAAAA 1560
Qy 1408 GGAAGAGGACAGAAACAGATTGAACCAACACCACTCATATGACCAAAATGAGAACATT 1467
Db 1561 TGCAGTCTCTGGCCCNAGATAGACCCGACTCCCGTGATGCTCATCTCCAAACCAACATG 1620
Qy 1468 ACAGTGTGTAGAATAGGCTCGGAAGAAAGACCAAGACACACTCAACCAATCAGAGACAGA 1527
Db 1621 TGGCCCGTGTATTGACGGGAACCTCAACGACCTTCGAACACAGCAGCGGTTGCAAGACCGG 1680
Qy 1528 ATGCTTAACTTTCATCAACATACCTTGCCTGGTGACTTTGGTTGGTTGACAAAAT 1587
Db 1681 ATGTTCAAAATTTGAACCTCACCCCGCTCTGGATCATGACTTTGGGAAGGTCAACCAAGCAG 1740
Qy 1588 GAA 1590
Db 1741 GAA 1743

Search completed: January 19, 2005, 17:05:44
Job time : 171.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2005, 11:50:09 ; Search time 1072 Seconds
(without alignments)
10821.780 Million cell updates/sec

Title: US-10-069-056-4
Perfect score: 2019
Sequence: 1 atggctggaagtcttactc.....agccgtgaactggactaa 2019

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1771	87.7	5121	US-10-647-111A-1	Sequence 1, Appli
2	943	46.7	5075	US-10-361-002-3	Sequence 3, Appli
3	943	46.7	5075	US-10-361-004-3	Sequence 3, Appli
4	307.4	15.2	374	US-10-466-894-1276	Sequence 1276, Ap
5	285.8	14.2	421	US-10-466-894-430	Sequence 430, App
6	284.8	14.1	423	US-10-466-894-375	Sequence 375, App
7	284.8	14.1	464	US-10-466-894-370	Sequence 370, App
8	284.8	14.1	473	US-10-466-894-367	Sequence 367, App
9	284.8	14.1	486	US-10-466-894-358	Sequence 358, App
10	284.8	14.1	491	US-10-466-894-428	Sequence 428, App
11	284.2	14.1	420	US-10-466-894-438	Sequence 438, App
12	283.8	14.1	343	US-10-466-894-388	Sequence 388, App

13	283.8	14.1	343	18	US-10-466-894-390	Sequence 390, App
14	283.8	14.1	403	18	US-10-466-894-433	Sequence 433, App
15	283.2	14.0	497	18	US-10-466-894-369	Sequence 369, App
16	280.4	13.9	340	18	US-10-466-894-383	Sequence 383, App
17	278.2	13.8	342	18	US-10-466-894-416	Sequence 416, App
18	273.8	13.6	420	18	US-10-466-894-432	Sequence 432, App
19	269.6	13.4	324	18	US-10-466-894-391	Sequence 391, App
20	263.4	13.0	339	18	US-10-466-894-1275	Sequence 1275, Ap
21	261.6	13.0	424	18	US-10-466-894-429	Sequence 429, App
22	163.8	8.1	468	18	US-10-466-894-524	Sequence 524, App
23	162.2	8.0	451	18	US-10-466-894-519	Sequence 519, App
24	162.2	8.0	457	18	US-10-466-894-518	Sequence 518, App
25	162.2	8.0	465	18	US-10-466-894-525	Sequence 525, App
26	162.2	8.0	469	18	US-10-466-894-520	Sequence 520, App
27	162.2	8.0	472	18	US-10-466-894-531	Sequence 531, App
28	162.2	8.0	515	18	US-10-466-894-444	Sequence 444, App
29	162.2	8.0	516	18	US-10-466-894-445	Sequence 445, App
30	162.2	8.0	516	18	US-10-466-894-451	Sequence 451, App
31	162.2	8.0	516	18	US-10-466-894-458	Sequence 458, App
32	162.2	8.0	517	18	US-10-466-894-462	Sequence 462, App
33	162.2	8.0	530	18	US-10-466-894-442	Sequence 442, App
34	162.2	8.0	530	18	US-10-466-894-450	Sequence 450, App
35	161	8.0	516	18	US-10-466-894-447	Sequence 447, App
36	160.6	8.0	530	18	US-10-466-894-456	Sequence 456, App
37	159	7.9	468	18	US-10-466-894-523	Sequence 523, App
38	133.4	6.6	247	18	US-10-466-894-471	Sequence 471, App
39	115.4	5.7	1386	9	US-09-792-630-32	Sequence 32, Appl
40	115.4	5.7	1386	10	US-09-953-351-32	Sequence 32, Appl
41	115.4	5.7	1386	13	US-10-080-376-32	Sequence 38, Appl
42	115.4	5.7	1386	14	US-10-082-671-38	Sequence 32, Appl
43	115.4	5.7	1386	14	US-10-097-100-32	Sequence 32, Appl
44	115.4	5.7	1386	15	US-10-023-208-32	Sequence 28, Appl
45	113.8	5.6	1884	9	US-09-792-630-28	

ALIGNMENTS

RESULT 1
US-10-647-111A-1
; Sequence 1, Application US/10647111A
; Publication No. US20040209240A1
; GENERAL INFORMATION:
; APPLICANT: IGGO, RICHARD
; APPLICANT: MALERBA, MADDALENA
; TITLE OF INVENTION: ANTI-NEOPLASTIC VIRUSES
; FILE REFERENCE: 604-691
; CURRENT APPLICATION NUMBER: US/10/647,111A
; CURRENT FILING DATE: 2003-08-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 1
; LENGTH: 5121
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Parvovirus H1
; OTHER INFORMATION: with promoter P4 and left hairpin from WVM (pHH1)
US-10-647-111A-1

Query Match	87.7%	Score	1771	DB	18	Length	5121
Best Local Similarity	92.3%	Pred. No.	0				
Matches	1864	Conservative	0	Mismatches	155	Indels	0
Gaps	0						
QY	1	ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA	60				
Db	282	ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA	341				
QY	61	AGTAAACGAGAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGAAAA	120				
Db	342	AGTAAACGAGAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGAAAA	401				
QY	121	GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGCAGCTGAAATCTTTACAA	180				

Db 402 GATATCGGATGGAATAGTTACAAAAGAGCTCGAGGAGCAGAGCTGAATCTTTACAA 461
QY 181 CGAGGAGCGAAATCTCTGGGACCAAGAGGACATGGAATGGGAAACCAAGTGGAT 240
Db 462 CGAGGAGCGGAAATCTCTGGGACCAAGAGGACATGGAATGGGAAACCAAGTGGAT 521
QY 241 GAAATGACCAAAAGCAAGATTAATCTTTTGAATCTTTGGTGTAAATAATGTTTATTTGAA 300
Db 522 GAAATGACCAAAAGCAAGATTAATCTTTTGAATCTTTGGTGTAAATAATGTTTATTTGAA 581
QY 301 GTGCTTAACAAAGAAATATATTTCTGTGTGATGTTAATGTTTGTGCAACATGAATGG 360
Db 582 GTGCTTTAAACAAAGAAATATATTTCTGTGTGATGTTAATGTTTGTGCAACATGAATGG 641
QY 361 GGAAGAGACCAAGCTGGCACTGCCATGTACTAATTTGGAGGAAGGACATTTAGTCAAGCT 420
Db 642 GGAAGAGACCAAGCTGGCACTGCCATGTACTAATTTGGAGGAAGGACATTTAGTCAAGCT 701
QY 421 CAAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCC 480
Db 702 CAAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCC 761
QY 481 TGTAAATGTGCAACTAACACAGCTGAAAGAAATTAATACTAAGAGAAATAGCAGAGCAAT 540
Db 762 TGTAAATGTGCAACTAACACAGCTGAAAGAAATTAATACTAAGAGAAATAGCAGAGCAAT 821
QY 541 GAGTGGTGTACTCTACTTACTTATAGCATAAGCAACCAAAAGACATATACCAAGTGT 600
Db 822 GAGTGGTGTACTCTACTTACTTATAGCATAAGCAACCAAAAGACATATACCAAGTGT 881
QY 601 GTTCTTTTGGAAACATGATGTTGCTTACTATTTTTTAACTAAAGAAAAATAAGCACTAGT 660
Db 882 GTTCTTTTGGAAACATGATGTTGCTTACTATTTTTTAACTAAAGAAAAATAAGCACTAGT 941
QY 661 CCACCAAGAGACGGAGCTATTTTCTTAGCAGTACTCTGGCTGGAAACTAACTTTTAA 720
Db 942 CCACCAAGAGACGGAGCTATTTTCTTAGCAGTACTCTGGCTGGAAACTAACTTTTAA 1001
QY 721 AAAGAGGCGGCGCCATCTAGTGAGCAAACTATACACTGATCAGATCGGCGCAGAAACG 780
Db 1002 AAAGAGGCGGCGCCATCTAGTGAGCAAACTATACACTGATCAGATCGGCGCAGAAACG 1061
QY 781 GTTGAACCAACAGTAACCACTGGCGGAGAACTAAGCGGCGGAGAAATCAAACTAAAAA 840
Db 1062 GTCGAGACCAAGTACCACTGACAGGAGCTAAGCGGCGGAGAAATCAAACTAGAGAG 1121
QY 841 GAAGTGTCTATTAATACTACATTAAGAGCTGGTGATTAAGAGTAACTTCACTCAGAG 900
Db 1122 GAGGTCTCGATTAAACCACTCAAGAGTGGTACATAAAAGAGTAACTTCACTCAGAG 1181
QY 901 GACTGGATGATGTCAGGCGAGACAGTTACATTTGAATGATGCTCAACAGGTGGAGAA 960
Db 1182 GACTGGATGATGTCAGGCGAGACAGTTACATTTGAATGATGCTCAACAGGTGGAGAA 1241
QY 961 AACCTGTGAAAAATACGCTAGAGATTTGTACACTAACTTAGCCAGAAACCAAAACAGCA 1020
Db 1242 AACTGTCTTAAAAATACACTAGAGATCTGTACACTGACTCTAGCAGAAACCAAAACAGCC 1301
QY 1021 TTTGACTTAATTTAGAAAAGCTGAAACCAAGCAACCACTAACCACTTTTCACTGGCTGAC 1080
Db 1302 TTTGACTTGAATTTGGAAAAAGCTGAAACCAAGCAACCACTAACCACTTTTCCATGGCTAGC 1361
QY 1081 ACAAGAACCTGCAAGATTTTGTGCTTTTCTGCTGGAATCTATGTTAAAGTTTGGCAATGCT 1140
Db 1362 ACCAGAACCTGTAGATCTTTGCTGAGCATGGCTGGACATATATTAAGTCTGCCATGCC 1421
QY 1141 ATTTGTGTGTTTTTAAACAGACAGGAGGCAAAAGAAATCTGTTTTTATTTATGACCA 1200
Db 1422 ATCTGTGTGCTGAATACAGAGGAGGCAAAAGGAAACACTGTGCTCTTTTCCAGGACCA 1481
QY 1201 GCCAGCAGGCAATCTATTTATGCAAGCCATAGCACAAGCAGTTGGCAATGTTGTT 1260

Db 1482 GCAGACAGGCAAAATCTATTATTGCAACCCATAGCAACAGCAGTTGGTAAATGTTGTT 1541
QY 1261 TGTATAATGCAAGCAATGTAACATTTTCCATTTAATGACTGTACCAACAAAGAACTTGAT 1320
Db 1542 TGTACAATGCTGCAATGTAACATTTTCCATTTAATGACTGTACCAACAAAGAACTTGAT 1601
QY 1321 TGGGTGAAGAAGCTGTTAACTTTTGGACAGCAAGTAACACCAATTTAAAGCCATTTGCTCT 1380
Db 1602 TGGGTGAAGAAGCTGTTAACTTTTGGCCAGCAAGTAACCAATTTCAAGCTATTTGTTCT 1661
QY 1381 GGTCAAACTATTTCGATGATCAAAAGGAGGAGCAAGCAAGATTCGAACCAACCA 1440
Db 1662 GGCACCAACATAGCATGATCAAAAGGAGGAGGAGCAAGATTCGAACCAACCA 1721
QY 1441 GTCATCATGACCAACCAATGAGAACATTCAGTGGTCAGAAATAGGCTGCGAAGAAAGCA 1500
Db 1722 GTTATTATGACCAACCAACGAGAACATTCAGTGGTGTAGAAATAGGCTGTGAGGAAGCA 1781
QY 1501 GAAACACATCAACCAATCAGACAGAAATGCTTTAACTTCACTCATCAACATACCTTGCCT 1560
Db 1782 GAAACACATCAACCAATCAGACAGAAATGCTTCAACATTCACCTGACACGTACACTACCT 1841
QY 1561 GGTGACTTTGTTTGGTTGACAAAAATGAATGCCCCATGATTTGCTTGGTTGGTAAAG 1620
Db 1842 GGTGACTTTGTTTGGTTGATAGACAGAAATGGCCCTCTGATCTGTGCTTGGTTGGTGAAG 1901
QY 1621 AATGTTTACCAATCTACATGCAAGCTACTGTGCTTAAATGGGGCAAGTTCTGTATGG 1680
Db 1902 AATGTTTACCAATCTACATGCTTGTACTGTCTTAAATGGGGCAAGTTCTGTATGG 1961
QY 1681 TCAGAAATCTGGGCGGAGCAAAAGTGGCACTCTCTATATAAATTTACTAGTTTGGCAAGC 1740
Db 1962 TCAGAGGACTGGGCGGAGCGGAGCTAGACCTCTTATAAATTCGCTAGTTTCAATGCGC 2021
QY 1741 TCACCAATTCAGCACACCGAAAGTAGCCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800
Db 2022 TCACCAATCTGATCTCGAGAAAGTGGCTCTCAGCCAAACTTACGCTTACTCCACTT 2081
QY 1801 GCATCGGATCTCGAGACCTGGCTTTTAGAGCTTTGGAGCAGACCAAAATCTCTGTGGG 1860
Db 2082 GCATCGGACCTTTGCGGACCTAGCTCTAGAGCTTTGGAGCAGACCAAAATCTCTGTGGG 2141
QY 1861 GGCATGCGAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCCTGCAAGATGCTCAA 1920
Db 2142 GGCATGCGAGCAAGCCAAACACTGGGGAAGCTGGTTCCACAGCTTCCCAAGGTGCTCAA 2201
QY 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTGGTGGCAACCG 1980
Db 2202 CGGAGCCCAACTTGGTCAGAGATCGAGGCGGATTTGAGAGCTTGTCTTCAAGAACAG 2261
QY 1981 TTGAAGAAAGACTTCAGCAGCGCTGAACTTGGACTAA 2019
Db 2262 TTGGAGCGGACTTCAACGAGGAGCTGACCTTGGACTAA 2300

RESULT 2

US-10-361-002-3
; Sequence 3, Application US/10361002
; Publication No. US20040170954A1
; GENERAL INFORMATION:
; APPLICANT: Clearant, Inc.
; APPLICANT: McKenney, Keith
; APPLICANT: Gillmeister, Lidja
; APPLICANT: Marlowe, Kristina
; APPLICANT: Armistead, David
; TITLE OF INVENTION: Pathogen Inactivation Assay
; FILE REFERENCE: CI-0043
; CURRENT APPLICATION NUMBER: US/10/361,002
; CURRENT FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 5075

[illegible]

```

RESULT 3
US-10-36
; Sequen
; Public
; GENERA
; APPLI
; APPLI
; APPLI
; APPLI
; APPLI
; TITLE
; FILE
; CURRE

```



```
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; TITLE OF INVENTION: AND THEIR USE AS MEDICINES
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1276
; LENGTH: 374
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 291, 300, 310, 316, 347, 348, 370, 371, 374
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-1276

Query Match      15.2%; Score 307.4; DB 18; Length 374;
Best Local Similarity 88.2%; Pred. No. 1.7e-70;
Matches 329; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 174 TTTTACACGAGGCGGAACTACTTTGGGACCAAGGAGGACATGGAATGGGAACCCAC 233
Db 1 TTTTACACGAGGCGGAGACCTTTGGGACCAAGGAGGACATGGAATGGGAGAGCGC 60

QY 234 AGTGGATGAAATGACCAAAAGCAAGTATTCTTTTGTATCTTTTGTAAAAAATGTTT 293
Db 61 AGTGGATGACATGACCAAAAGCAAGTATTCTTTTGTATCTTTTGTAAAGAGTGT 120

QY 294 ATTTGAAGTCTTAAACAAAGATATATTTCTTGTGTATGTTAAATGTTGTGAACA 353
Db 121 GTTTGAAGTCTCAGCACAAAGAACATAGCTCTCTAGTAATGTTTACTTTGTCGAGCA 180

QY 354 TGAATGGGAAAGACCAAGGCTGGCAGTCCCTACTTAATTTGGAGGAAGCACTTTAG 413
Db 181 TGAATGGGAAAGACCAAGGCTGGCAGTCTGCTGATTTGGAGGCAAGCACTTTAG 240

QY 414 TCAAGCTCAAGGAAATGGTGGAGAGCAACTAAATGTTTACTGGAGCAGATGGTTGGT 473
Db 241 TCAACCTCAAGGAAATGGTGGAGAGGAGCTAAATGTTGCTAGGTANATGGTTGGN 300

QY 474 AACAGCTGTAAATGTCGAATACACCAAGCTGAAAGAAATTAATACTAAGAGAAATAGCAGA 533
Db 301 GACTGCTGNAATGTNCAACTAACACCAAGCTGAAAGAAATTAATACTGNNAGAAATAGCAGA 360

QY 534 AGACAATGAGTGG 546
Db 361 GGACAGTGANNNG 373

RESULT 5
US-10-466-894-430
; Sequence 430, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; TITLE OF INVENTION: AND THEIR USE AS MEDICINES
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1276
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 422
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-375
```

```
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 430
; LENGTH: 421
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 276, 405, 417, 419, 421
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-430

Query Match      14.2%; Score 285.8; DB 18; Length 421;
Best Local Similarity 87.6%; Pred. No. 9.4e-65;
Matches 311; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGCGGAACTACTTTGGGACCAAGGAGGACATGGAATGGGAACCA 232
Db 59 CTTTACACGAGGCGGAGACCTTTGGGACCAAGGAGGACATGGAATGGGAGAGCG 118

QY 233 CAGTGTGAAATGACCAAAAGCAAGTATTCTTTTGTATCTTTTGTAAAAAATGTT 292
Db 119 CAGTGTGACATGACCAAAAGCAAGTATTCTTTTGTATCTTTTGTAAAGAGTGT 178

QY 293 TATTTGAAAGTCTTAAACAAAGATATATTTCTGTGTATGTTAAATGTTGTGCAAC 352
Db 179 TGTTTGAAAGTCTCAGCACAAAGACATAGCTCTTAGTATGTTACTTTGTCGAGC 238

QY 353 ATGAATGGGAAAGCAAGGCTGGCAGTCCATGTAATAATGGAGAAAGACTTTTA 412
Db 239 ATGAATGGGAAAGCAAGGCTGGCAGTCTGATGCTGCTGATGTTGGAGGCAAGACTTTA 298

QY 413 GTCAAGCTCAAGGAAATGGTGGAGAGCAACTAAATGTTTACTTGAGCAGATGGTTGG 472
Db 299 GTCAAGCTCAAGGAAATGGTGGAGAGGAGCAGCTAAATGTTGACTGAGTAGATGGTTGG 358

QY 473 TAAACGCTGTATGTGCAACTAACACCAAGCTGAAAGAAATTAATACTAAGAGAAAT 527
Db 359 TGAAGCTGTATGTTCAACTAACACCAAGCTGAAAGAAATTAATACTAAGAGAAAT 413

RESULT 6
US-10-466-894-375
; Sequence 375, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; TITLE OF INVENTION: AND THEIR USE AS MEDICINES
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 375
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 422
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-375
```



```
Db 372 TGACTGCTGTAATGTTCAACTAACACCACTGAAACCAAAAAA 415
|||||
RESULT 9
US-10-466-894-358
; Sequence 358, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 358
; LENGTH: 486
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 413, 428, 435, 448, 457, 462
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-358

Query Match 14.1%; Score 284.8; DB 18; Length 486;
Best Local Similarity 89.2%; Pred. No. 1.9e-64;
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGAGCGGAACTACTTTGGACCAAAAGCGAGGACATGGAAATGGGAAACCA 232
Db 54 CTTTACACGAGGAGCGGAGACCACTTTGGACCAAAAGCGAGGACATGGAAATGGGAGCG 113
QY 233 CAGTGAATGAATGACCAAAAGCAAGTATTCTTTTGTATCTTTTGGTTAAAAATGTT 292
Db 114 CAGTGAATGACATGACCAAAAGCAAGTATTCTTTTGTATCTTTTGGTTAAAAATGTT 173
QY 293 TATTGAAGTCTTAACACAAAGATATATTTCTGCTGATGTTAATTTGTTTGTGCAAC 352
Db 174 TGTTGAAGTCTCAGCACAAAGACATAGCTCTCTAGTAATGTTACTTGTGCGAGC 233
QY 353 ATGAATGGGAAAAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGACTTTA 412
Db 234 ATGAATGGGAAAAGACCAAGGCTGGCACTGTCTGCTGATGTTAGGAGGCAAGACTTTA 293
QY 413 GTCAAGCTCAAGGAAAATGTTGAGAGGCAACTAAATGTTTACTGGAGCAGATGTTGG 472
Db 294 GTCAAGCTCAAGGAAAATGTTGAGAGGCAAGTAAATGTTGTTACTGGAGTATGTTGG 353
QY 473 TAACAGCTGTATGTGCACTTAACACCACTGAAACCAATTTAA 516
Db 354 TGACTGCTGTAATGTTCAACTAACACCACTGAAACCAAAAAA 397

RESULT 10
US-10-466-894-428
; Sequence 428, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 358
; LENGTH: 486
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 413, 428, 435, 448, 457, 462
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-428

Query Match 14.1%; Score 284.8; DB 18; Length 491;
Best Local Similarity 89.2%; Pred. No. 1.9e-64;
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGAGCGGAACTACTTTGGACCAAAAGCGAGGACATGGAAATGGGAAACCA 232
Db 62 CTTTACACGAGGAGCGGAGACCACTTTGGACCAAAAGCGAGGACATGGAAATGGGAGCG 121
QY 233 CAGTGAATGAATGACCAAAAGCAAGTATTCTTTTGTATCTTTTGGTTAAAAATGTT 292
Db 122 CAGTGAATGACATGACCAAAAGCAAGTATTCTTTTGTATCTTTTGGTTAAAAATGTT 181
QY 293 TATTGAAGTCTTAACACAAAGATATATTTCTGCTGATGTTAATTTGTTTGTGCAAC 352
Db 182 TGTTGAAGTCTCAGCACAAAGACATAGCTCTCTAGTAATGTTACTTGTGCGAGC 241
QY 353 ATGAATGGGAAAAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGACTTTA 412
Db 242 ATGAATGGGAAAAGACCAAGGCTGGCACTGTCTGCTGATGTTAGGAGGCAAGACTTTA 301
QY 413 GTCAAGCTCAAGGAAAATGTTGAGAGGCAACTAAATGTTTACTGGAGCAGATGTTGG 472
Db 302 GTCAAGCTCAAGGAAAATGTTGAGAGGCAAGTAAATGTTGTTACTGGAGTATGTTGG 361
QY 473 TAACAGCTGTATGTGCACTTAACACCACTGAAACCAATTTAA 516
Db 362 TGACTGCTGTAATGTTCAACTAACACCACTGAAACCAAAAAA 405

RESULT 11
US-10-466-894-438
; Sequence 438, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 438
```

```
/ LENGTH: 420
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 418
/ OTHER INFORMATION: n = A,T,C or G
US-10-466-894-438

Query Match
Best Local Similarity 14.1%; Score 284.2; DB 18; Length 420;
Matches 310; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 173 CTTTCAACGAGGAGCGGAACTACTTGGACCAAAAGCGAGGACATGGAATGGGAACCA 232
Db 64 CTTTCAACGAGGAGCGGAGCACCTTGGACCAAAAGCGAGGACATGGAATGGGAGCG 123

QY 233 CAGTGAATGAATGACCAAAAGCAAGTATTCATTTTGAATCTTTGGTTTAAAAATGTT 292
Db 124 CAGTGAATGACATGACCAAAAGCAAGTATTTATTTTGAATCTTTGGCTAAGAAGTGT 183

QY 293 TATTTCAAGTCTTAACACAAAGAAATATATTTCTCGTGATGTTAATTTGTTGCAAC 352
Db 184 TGTTTGAAGTCTCAGACAAAGAAATAGCTCTAGTAAGTTACTTTGTTTGCAGC 243

QY 353 ATGAATGGGAAAGACCAAGGCTGGCAGCTGCCATGCTACTTAATTTGGAGAAAGCACTTTA 412
Db 244 ATGAATGGGAAAGACCAAGGCTGGCAGCTGCCATGCTACTTAATTTGGAGCAAGGACTTTA 303

QY 413 GTCAGCTCAAGGAAATGTTGAGAAAGCACTAAATGTTTACTGGAGCAGATGTTGG 472
Db 304 GTCAACCTCAAGGAAATGTTGAGAAAGCACTAAATGTTTACTGGAGTAGATGTTGG 363

QY 473 TACAGCTGTAATGTGCAACTAACACCAAGCTGAAAGAAATTAACCTAAGAGAA 525
Db 364 TGACTGCTGTAATGTTCACCTAACACCAAGCTGAAAGAAATTAACCTAAGAGAA 416

RESULT 12
US-10-466-894-388
/ Sequence 388, Application US/10466894
/ Publication No. US20040241671A1
/ GENERAL INFORMATION:
/ APPLICANT: Telerman, Adam
/ APPLICANT: Amson, Robert
/ APPLICANT: Tuijnder, Marius
/ APPLICANT: Subini, Laurent
/ TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
/ TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
/ FILE REFERENCE: 10918-014-999
/ CURRENT APPLICATION NUMBER: US/10/466,894
/ PRIOR APPLICATION NUMBER: PCT/FR02/00273
/ PRIOR FILING DATE: 2002-01-23
/ PRIOR APPLICATION NUMBER: FR01/00899
/ PRIOR FILING DATE: 2001-01-23
/ NUMBER OF SEQ ID NOS: 2270
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 388
/ LENGTH: 343
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
US-10-466-894-388

Query Match
Best Local Similarity 14.1%; Score 283.8; DB 18; Length 343;
Matches 306; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 174 TTTTCAACGAGGAGCGGAACTACTTGGACCAAAAGCGAGGACATGGAATGGGAACCA 233
Db 1 TTTTCAACGAGGAGCGGAGCACCTTGGACCAAAAGCGAGGACATGGAATGGGAGCGC 60

QY 234 AGTGAATGAATGACCAAAAGCAAGTATTCATTTTGAATCTTTGTTTAAAAATGTTT 293
Db 61 AGTGAATGACATGACCAAAAGCAAGTATTTATTTTGAATCTTTGTTTAAAAAGTGT 120

QY 294 ATTTGAAGTCTTAACACAAAGAAATATTTCTCGTGATGTTAATTTGTTTGAACAA 353
Db 121 GTTTGAAGTCTCAGACAAAGAAATAGCTCTAGTAAGTTACTTTGTTTGCAGCA 180

QY 354 TGAATGGGAAAGACCAAGGCTGGCAGCTGCCATGCTACTTAATTTGGAGAAAGCACTTTAG 413
Db 181 TGAATGGGAAAGACCAAGGCTGGCAGCTGCCATGCTACTTAATTTGGAGCAAGGACTTTAG 240

QY 414 TCAAGCTCAAGGAAATGTTGAGAAAGCACTAAATGTTTACTGGAGCAGATGTTGGT 473
Db 241 TCAACCTCAAGGAAATGTTGAGAAAGCACTAAATGTTTACTGGAGTAGATGTTGGT 300

QY 474 AACAGCTGTAATGTGCAACTAACACCAAGCTGAAAGAAATTAAC 516
```

```
234 ACTGATGAATGACCAAAAGCAAGTATTCATTTTGAATCTTTGGTTTAAAAATGTTT 293
61 ACTGATGACATGACCAAAAGCAAGTATTTATTTTGAATCTTTGGTTTAAAGAGTGT 120

294 ATTTGAAGTCTTAACACAAAGAAATATTTCTCGTGATGTTAATTTGTTTGAACAA 353
121 GTTTGAAGTCTCAGACAAAGAAATAGCTCTAGTAAGTTACTTTGTTTGCAGCA 180

354 TGAATGGGAAAGACCAAGGCTGGCAGCTGCCATGCTACTTAATTTGGAGAAAGCACTTTAG 413
181 TGAATGGGAAAGACCAAGGCTGGCAGCTGCCATGCTACTTAATTTGGAGCAAGGACTTTAG 240

414 TCAAGCTCAAGGAAATGTTGAGAAAGCACTAAATGTTTACTGGAGCAGATGTTGGT 473
241 TCAACCTCAAGGAAATGTTGAGAAAGCACTAAATGTTTACTGGAGTAGATGTTGGT 300

474 AACAGCTGTAATGTGCAACTAACACCAAGCTGAAAGAAATTAAC 516
301 GACTGCTGTAATGTTCACCTAACACCAAGCTGAAAGAAATTAAC 343

RESULT 13
US-10-466-894-390
/ Sequence 390, Application US/10466894
/ Publication No. US20040241671A1
/ GENERAL INFORMATION:
/ APPLICANT: Telerman, Adam
/ APPLICANT: Amson, Robert
/ APPLICANT: Tuijnder, Marius
/ APPLICANT: Subini, Laurent
/ TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
/ TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
/ FILE REFERENCE: 10918-014-999
/ CURRENT APPLICATION NUMBER: US/10/466,894
/ PRIOR APPLICATION NUMBER: PCT/FR02/00273
/ PRIOR FILING DATE: 2002-01-23
/ PRIOR APPLICATION NUMBER: FR01/00899
/ PRIOR FILING DATE: 2001-01-23
/ NUMBER OF SEQ ID NOS: 2270
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 390
/ LENGTH: 343
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
US-10-466-894-390

Query Match
Best Local Similarity 14.1%; Score 283.8; DB 18; Length 343;
Matches 306; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 174 TTTTCAACGAGGAGCGGAACTACTTGGACCAAAAGCGAGGACATGGAATGGGAACCA 233
Db 1 TTTTCAACGAGGAGCGGAGCACCTTGGACCAAAAGCGAGGACATGGAATGGGAGCGC 60

QY 234 AGTGAATGAATGACCAAAAGCAAGTATTCATTTTGAATCTTTGTTTAAAAATGTTT 293
Db 61 AGTGAATGACATGACCAAAAGCAAGTATTTATTTTGAATCTTTGTTTAAAAAGTGT 120

QY 294 ATTTGAAGTCTTAACACAAAGAAATATTTCTCGTGATGTTAATTTGTTTGAACAA 353
Db 121 GTTTGAAGTCTCAGACAAAGAAATAGCTCTAGTAAGTTACTTTGTTTGCAGCA 180

QY 354 TGAATGGGAAAGACCAAGGCTGGCAGCTGCCATGCTACTTAATTTGGAGAAAGCACTTTAG 413
Db 181 TGAATGGGAAAGACCAAGGCTGGCAGCTGCCATGCTACTTAATTTGGAGCAAGGACTTTAG 240

QY 414 TCAAGCTCAAGGAAATGTTGAGAAAGCACTAAATGTTTACTGGAGCAGATGTTGGT 473
Db 241 TCAACCTCAAGGAAATGTTGAGAAAGCACTAAATGTTTACTGGAGTAGATGTTGGT 300

QY 474 AACAGCTGTAATGTGCAACTAACACCAAGCTGAAAGAAATTAAC 516
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:18:38 ; Search time 99 Seconds
(without alignments)
2435.011 Million cell updates/sec

Title: US-10-069-056-6
Perfect score: 3638
Sequence: 1 MAGNAYSDEVLGNTWLKKE.....RACFGAEPLKDFSEPLNLD 672

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : A_Geneseq_23Sep04:*
- 1: Geneseqp1980s:*
 - 2: Geneseqp1990s:*
 - 3: Geneseqp2000s:*
 - 4: Geneseqp2001s:*
 - 5: Geneseqp2002s:*
 - 6: Geneseqp2003as:*
 - 7: Geneseqp2003bs:*
 - 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3628	99.7	672	4	AAU72704 Parvoviru
2	3625	99.6	672	4	AAU72702 Parvoviru
3	3620	99.5	672	4	AAU72708 Parvoviru
4	3620	99.5	672	4	AAU72710 Parvoviru
5	3620	99.5	672	4	AAU72706 Parvoviru
6	2373	65.2	690	1	AAU40306 Sequence
7	455.5	12.5	550	3	AAU58164 Adeno ass
8	455.5	12.5	550	3	AAU11409 Adeno-ass
9	455.5	12.5	610	3	AAU58159 Adeno ass
10	455.5	12.5	610	4	AAU97720 Rep prote
11	455.5	12.5	610	5	AAU11404 Adeno-ass
12	455.5	12.5	610	5	AAU22887 Adeno-ass
13	455.5	12.5	610	5	AAU28644 Adeno-ass
14	455.5	12.5	610	5	AAU26940 Adeno ass
15	455.5	12.5	610	6	AAU64865 Rep prote
16	455.5	12.5	610	6	AAU64760 Adeno ass
17	455.5	12.5	610	6	AAU43398 Adeno-ass
18	455.5	12.5	610	7	AAU40280 Adeno-ass
19	455.5	12.5	610	8	AAU58895 Predeterm
20	452.5	12.4	626	2	AAU85384 Barbaty d
21	452.5	12.4	626	4	AAU97724 NSI prote
22	452.5	12.4	626	5	AAU22891 Muscovy d
23	452.5	12.4	626	5	AAU28648 Muscovy d
24	452.5	12.4	626	5	AAU26944 Muscovy d
25	452.5	12.4	626	6	AAU64869 NSI prote

26	452.5	12.4	626	6	AAU64764 Parvoviru
27	452.5	12.4	626	6	AAU43402 Muscovy d
28	452.5	12.4	626	7	AAU40288 Muscovy d
29	452	12.4	627	4	AAU97722 Rep prote
30	452	12.4	627	5	AAU22889 Barbatie
31	452	12.4	627	5	AAU28646 Barbatie
32	452	12.4	627	5	AAU26942 Barbatie
33	452	12.4	627	6	AAU64867 Rep prote
34	452	12.4	627	6	AAU64762 Parvoviru
35	452	12.4	627	6	AAU43400 Barbatie
36	452	12.4	627	7	AAU40284 Barbatie
37	445.5	12.2	625	7	AAU62760 Adeno ass
38	437.5	12.0	536	2	AAU46311 AAU4 Rep
39	437.5	12.0	536	6	AAU73938 Adeno-ass
40	437.5	12.0	623	2	AAU46307 AAU4 Rep
41	437.5	12.0	623	4	AAU46312 AAU4 Rep
42	437.5	12.0	623	4	AAU97712 Rep prote
43	437.5	12.0	623	5	AAU28636 Adeno-ass
44	437.5	12.0	623	5	AAU26932 Adeno ass
45	437.5	12.0	623	6	AAU64857 Rep prote

ALIGNMENTS

RESULT 1
AAU72704
ID AAU72704 standard; protein; 672 AA.
XX
AC AAU72704;
XX
DT 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
DE Parvovirus non-structure protein 1 (NS1) variant (S283A).
XX
KW NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoural disease; gene therapy; mutant; mutein; variant.
XX
OS Parvovirus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 283 /note= "Wild type Ser substituted with Ala"
XX
PN EP1077260-AL.
XX
PD 21-FEB-2001.
XX
PF 13-AUG-1999; 99EP-00115161.
XX
PR 13-AUG-1999; 99EP-00115161.
XX
PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
PI Nueesch J, Rommelaere J;
XX WPI: 2001-212717/22.
XX N-PSDB; AAD02799.
DR Novel parvovirus non-structure protein variant, useful for treating
DR tumoral diseases, has a shifted equilibrium between DNA replication and
XX transcription activities, and cytotoxic activity.
XX Claim 6; Page 14-16; 41pp; English.
XX
CC The present sequence is parvovirus non-structure protein 1 (NS1) variant
CC (S283A). The invention relates to the variants of the parvovirus non-
CC structure protein (NS1) having a shifted equilibrium between the DNA
CC replication and transcription activities, and the cytotoxicity activity.
CC These variants are useful as vectors for treating tumoural diseases. The
CC variant DNAs are useful for gene therapy. (Updated on 06-AUG-


```
Db 541 NGYQSTWASCAKWKVPDMSNWAEPKVPPTPINLLGSARSPPTPKSTPLSQNYALTPL 600
Qy 601 ASDLEDLALPWSPTNPVAGTAETQNTGEGSKACQDQGLSPTWSEIEDLRACFGAEP 660
Db 601 ASDLEDLALPWSPTNPVAGTAETQNTGEGSKACQDQGLSPTWSEIEDLRACFGAEP 660
Qy 661 LKKDFSEPLNLD 672
Db 661 LKKDFSEPLNLD 672

RESULT 3
AAY72708
ID AAY72708 standard; protein; 672 AA.
XX AC AAY72708;
XX 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX Parvovirus non-structure protein 1 (NS1) variant (T394A).
DE NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoral disease; gene therapy; mutant; mutein; variant.
XX Parvovirus.
OS Synthetic.
XX FH Key Location/Qualifiers
FT Misc-difference 394 /note= "Wild type Thr substituted with Ala"
XX XX
XX EP1077260-A1.
PD 21-FEB-2001.
XX PF 13-AUG-1999; 99EP-00115161.
XX PR 13-AUG-1999; 99EP-00115161.
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX PI Nueesch J, Rommelaere J;
XX WP1; 2001-212717/22.
DR N-PSDB; AAD02803.
XX PT Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX PS Claim 6; Page 25-27; 41pp; English.
XX CC The present sequence is parvovirus non-structure protein 1 (NS1) variant
CC (T394A). The invention relates to the variants of the parvovirus non-
CC structure protein (NS1) having a shifted equilibrium between the DNA
CC replication and transcription activities, and the cytotoxicity activity.
CC These variants are useful as toxins for treating tumoral diseases. The
CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX SQ Sequence 672 AA;

Query Match 99.5%; Score 3620; DB 4; Length 672;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 669; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAGNAYDEVLGATNWLKESNQEVFVFKNENVQLNGKDIGNWSYKKELODELSLQ 60
Db 1 MAGNAYDEVLGATNWLKESNQEVFVFKNENVQLNGKDIGNWSYKKELODELSLQ 60
Qy 61 RGAETTWQSDMEWETTVDEMTHKQVIFDLSLVKKCLFEVLANTKNIFPGDVNWFQHEW 120
```

XX Nueesch J, Rommelaere J;
XX WPI; 2001-212717/22.
XX N-PSDB; AAD02805.
XX Novel parvovirus non-structure protein variant, useful for treating
XX tumoral diseases, has a shifted equilibrium between DNA replication and
XX transcription activities, and cytotoxic activity.
XX Claim 6; Page 30-32; 41pp; English.
XX The present sequence is parvovirus non-structure protein 1 (NS1) variant
XX (T463A). The invention relates to the variants of the parvovirus non-
XX structure protein (NS1) having a shifted equilibrium between the DNA
XX replication and transcription activities, and the cytotoxicity activity.
XX These variants are useful as toxins for treating tumoral diseases. The
XX variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
XX 2003 to correct OS field.)
XX SQ Sequence 672 AA;
Query Match 99.5%; Score 3620; DB 4; Length 672;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 669; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 MAGNAYSDEVLGATNWLKESQVFSFVKNNVQVNGKDIGWNSYKKELODELKSLQ 60
DB 1 MAGNAYSDEVLGATNWLKESQVFSFVKNNVQVNGKDIGWNSYKKELODELKSLQ 60
QY 61 RGAETTWQSEDMEWETTVDEMTEKKQVIFDLSLVKKCLFEVLNTKNIFFGDVNVFVQHEW 120
DB 61 RGAETTWQSEDMEWETTVDEMTEKKQVIFDLSLVKKCLFEVLNTKNIFFGDVNVFVQHEW 120
QY 121 GKQGWCHVLIIGKQDSOAGQKWRRLNYVSRWLVTACNVQLTPAERIKLREIAEDN 180
DB 121 GKQGWCHVLIIGKQDSOAGQKWRRLNYVSRWLVTACNVQLTPAERIKLREIAEDN 180
QY 181 EWTLLTYKHQTKQDKYTKCVLFGNMIAIYFLLTKKISTSPRDGGYFLSSDSGWKTNPL 240
DB 181 EWTLLTYKHQTKQDKYTKCVLFGNMIAIYFLLTKKISTSPRDGGYFLSSDSGWKTNPL 240
QY 241 KEGERHLVSKLYTDDMRPETVETTVTAQETKRGRIQTKKEVAIKTLLKELVHKRVTSPE 300
DB 241 KEGERHLVSKLYTDDMRPETVETTVTAQETKRGRIQTKKEVAIKTLLKELVHKRVTSPE 300
QY 301 DWMQPDSDYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKASTSKLTNFSLPD 360
DB 301 DWMQPDSDYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKASTSKLTNFSLPD 360
QY 361 TRTCRIFAFHGMNIVKYCHAIACCVLNRQGGKRNVLFGHPASTGKSIIAQAIAQAVGNVG 420
DB 361 TRTCRIFAFHGMNIVKYCHAIACCVLNRQGGKRNVLFGHPASTGKSIIAQAIAQAVGNVG 420
QY 421 CYNAAVNFNDCTKNLIWBEAGNFGQVQNFKAICSGQTIRIDQKGSQKIEPTP 480
DB 421 CYNAAVNFNDCTKNLIWBEAGNFGQVQNFKAICSGQTIRIDQKGSQKIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDPLGVKNEWPMICAWLYK 540
DB 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDPLGVKNEWPMICAWLYK 540
QY 541 NGYSTMASYCAKMGKVPDWSNABPKVPTPINLLGARSPTTTPKSTPLSONYALTEL 600
DB 541 NGYSTMASYCAKMGKVPDWSNABPKVPTPINLLGARSPTTTPKSTPLSONYALTEL 600
QY 601 ASDLEDLALEPWSPTNPVAGTAEHTONTGEAGSKACQDQQLSPTWSEIEDLRACFGAPP 660
DB 601 ASDLEDLALEPWSPTNPVAGTAEHTONTGEAGSKACQDQQLSPTWSEIEDLRACFGAPP 660
QY 661 LKXDFSEPLNLD 672
DB 661 LKXDFSEPLNLD 672

RESULT 5
AA72706
ID AAY72706 standard; protein; 672 AA.
XX
XX AAY72706;
XX AC AAY72706;
XX DT 06-AUG-2003 (revised)
XX DT 31-MAY-2001 (first entry)
XX
XX Parvovirus non-structure protein 1 (NS1) variant (T363A).
XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
XX tumoral disease; gene therapy; mutant; mutein; variant.
XX OS Parvovirus.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Misc-difference 363 /note= "Wild type Thr substituted with Ala"
XX PN EP1077260-A1.
XX PD 21-FEB-2001.
XX PF 13-AUG-1999; 99EP-00115161.
XX PR 13-AUG-1999; 99EP-00115161.
XX PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX PI Nueesch J, Rommelaere J;
XX WPI; 2001-212717/22.
XX DR N-PSDB; AAD02801.
XX PT Novel parvovirus non-structure protein variant, useful for treating
XX tumoral diseases, has a shifted equilibrium between DNA replication and
XX transcription activities, and cytotoxic activity.
XX PS Claim 6; Page 19-21; 41pp; English.
XX The present sequence is parvovirus non-structure protein 1 (NS1) variant
XX (T363A). The invention relates to the variants of the parvovirus non-
XX structure protein (NS1) having a shifted equilibrium between the DNA
XX replication and transcription activities, and the cytotoxicity activity.
XX These variants are useful as toxins for treating tumoral diseases. The
XX variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
XX 2003 to correct OS field.)
XX SQ Sequence 672 AA;
Query Match 99.5%; Score 3620; DB 4; Length 672;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 669; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 MAGNAYSDEVLGATNWLKESQVFSFVKNNVQVNGKDIGWNSYKKELODELKSLQ 60
DB 1 MAGNAYSDEVLGATNWLKESQVFSFVKNNVQVNGKDIGWNSYKKELODELKSLQ 60
QY 61 RGAETTWQSEDMEWETTVDEMTEKKQVIFDLSLVKKCLFEVLNTKNIFFGDVNVFVQHEW 120
DB 61 RGAETTWQSEDMEWETTVDEMTEKKQVIFDLSLVKKCLFEVLNTKNIFFGDVNVFVQHEW 120
QY 121 GKQGWCHVLIIGKQDSOAGQKWRRLNYVSRWLVTACNVQLTPAERIKLREIAEDN 180
DB 121 GKQGWCHVLIIGKQDSOAGQKWRRLNYVSRWLVTACNVQLTPAERIKLREIAEDN 180
QY 181 EWTLLTYKHQTKQDKYTKCVLFGNMIAIYFLLTKKISTSPRDGGYFLSSDSGWKTNPL 240
DB 181 EWTLLTYKHQTKQDKYTKCVLFGNMIAIYFLLTKKISTSPRDGGYFLSSDSGWKTNPL 240

QY 241 KGERHLVSKLYTDDMRPETVETTTTAQETKGRGRIOTKKEVAIKTTLKELVHKRVTSPE 300
 DB 241 KGERHLVSKLYTDDMRPETVETTTTAQETKGRGRIOTKKEVAIKTTLKELVHKRVTSPE 300
 QY 301 DWMMPQDSYIEMMAQPGGNNLKNLTLEICTLTARTKTAFTDLILEKAETSKLTNFSLPD 360
 DB 301 DWMMPQDSYIEMMAQPGGNNLKNLTLEICTLTARTKTAFTDLILEKAETSKLTNFSLPD 360
 QY 361 TRTRCIFAFAHGMNKKVCHCAICCVLNRQGGKRNITLPHGPASTGKSIIAQAIQAVGNVG 420
 DB 361 TRACRIFAFAHGMNKKVCHCAICCVLNRQGGKRNITLPHGPASTGKSIIAQAIQAVGNVG 420
 QY 421 CYNAAVNFPPNDCTNKNLIWVEEAGNFGQVNOFKKAIKCSGQIRIDQKGGSKQIEPTP 480
 DB 421 CYNAAVNFPPNDCTNKNLIWVEEAGNFGQVNOFKKAIKCSGQIRIDQKGGSKQIEPTP 480
 QY 481 VIMTTNENITVIRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNEMPMICAWLVK 540
 DB 481 VIMTTNENITVIRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNEMPMICAWLVK 540
 QY 541 NGYQSTMASYCAKWKGVDPDSNMAEPKVPPTPINLLGSARSPTTPKSTPLSQNYALTPL 600
 DB 541 NGYQSTMASYCAKWKGVDPDSNMAEPKVPPTPINLLGSARSPTTPKSTPLSQNYALTPL 600
 QY 601 ASDLEDLALPWSPTNPVAGTAETQNTGEGAGSKACODGQLSPTWSEIEDLRAACFCAEP 660
 DB 601 ASDLEDLALPWSPTNPVAGTAETQNTGEGAGSKACODGQLSPTWSEIEDLRAACFCAEP 660
 QY 661 LKXDFSEPLNLD 672
 DB 661 LKXDFSEPLNLD 672

RESULT 6
 AAP40306 standard; protein; 690 AA.
 AC AAP40306;
 DT 24-OCT-2003 (revised)
 DT 12-JAN-1992 (first entry)
 XX
 Sequence encoded by the double-stranded replicative form DNA of porcine parvovirus.
 KW Protein envelope; immunogen; vaccine; antigen; epitope.
 XX Porcine parvovirus; NADL-2 virulent strain.
 OS WO8402847-A.
 PN
 XX 02-AUG-1984.
 PD
 XX 19-JAN-1984; 84WO-US0000063.
 PF
 XX 19-JAN-1983; 83US-00459203.
 PR 06-JAN-1984; 84US-00567968.
 XX
 XX (AMGE-) AMGEN.
 PA
 XX
 XX Fox GW;
 PI
 XX
 XX WPI; 1984-201354/32.
 DR N-PSDB; AAN40252.
 XX
 Polypeptide obtd. by recombinant DNA methods - for vaccination against parvovirus infections in man and animals.
 PT
 XX
 PS Claim 6; Table II, Page 33-49; 80pp; English.
 XX
 XX The inventors claim an immunologically active polypeptide for the development of vaccinal immunity against parvovirus infection. Also

CC claimed are DNA sequences wholly or partly duplicative of defined sequences. The polypeptides are used in vaccines for conferring protection against parvovirus infections in man and animals. (Updated on 24-OCT-2003 to standardise OS field)

XX SQ Sequence 690 AA;

Query Match 65.2%; Score 2373; DB 1; Length 690;
 Best Local Similarity 67.7%; Pred. No. 1.7e-222;
 Matches 448; Conservative 71; Mismatches 127; Indels 16; Gaps 6;

QY 2 AGNAYSDVLTGATNMLKEKSNQEVFSFVKNNVQLNGKDIGNSYKKEIQEDELKSLQR 61
 DB 5 AGNTYSEEVLUKATNMLQDNAQKEAFSYVFKTKVNLNGKEIAWNNYKNTTDAEMINLQR 64
 QY 62 GAETTWQSDSEDMETTTVDKTKQVIFIDSLVKKCLFEVLNTKNTPPGVDNMFVQHEWG 121
 DB 65 GAETSDWQATDMESEIDSLTKRQVLIIDSLVKKCLFEGILQKLSPSDCYWFLOHEHG 124
 QY 122 KDQGWCHVULIGKDFSQAGKWRRLQNLVYWRWLVTACNVOLTTPAERIKLREIADNE 181
 DB 125 QDTGYHCHVLLGGGLQAGKWLRLQNLVYWRWLVTACNVOLTTPAERIKLREIADNE 184
 QY 182 WVTLLTYKHQTKQDYTKCVLFCNMLIAYFLTKKJSTSPRDGGYFLSDSGWKTNFKL 241
 DB 185 WSLLLTYTHKQTKQDYTKCVLFCNMLIAYFLTKKJSTSPRDGGYFLSDSGWKTNFKL 242
 QY 242 EGERHLVSKLYTDDMRPETVETTTTAQETKGRGRIOTKKEVAIKTTLKELVHKRVTSPE 301
 DB 243 EGERHLVSHLFTANKPETVETTTTAQETKGRGRIOTKKEVAIKTTLKELVHKRVTSPE 302
 QY 302 WMMQPDYSIEMMAQPGGNNLKNLTLEICTLTARTKTAFTDLILEKAETSKLTNFSLPD 361
 DB 303 WSMTPDYSIEMMAQPGGNNLKNLTLEICTLTARTKTAFTDLILEKAETSKLTNFSLPD 362
 QY 362 RTCRIFAFAHGMNKKVCHCAICCVL--NRQGGKENTVLFHGPASTGKSIIAQAIQAVGNV 419
 DB 363 RTCKIFSMHNNYIKVCHAITCVLKQTRRRKKYNSISCHGPASTGKSIIAQAIQAVGNV 422
 QY 420 GCYNAANVNFPPNDCTNKNLIWVEEAGNFGQVNOFKKAIKCSGQIRIDQKGGSKQIEPT 479
 DB 423 GCYNAANVNFPPNDCTNKNLIWVEEAGNFGQVNOFKKAIKCSGQIRIDQKGGSKQIEPT 482
 QY 480 PVIMTTNENITVIRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNEMPMICAWLV 539
 DB 483 PVIMTTNENITVIRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNEMPMICAWLV 542
 QY 540 KNGYQSTMASYCAKWKGVDPDSNMAEPKVPPTPINLLGSARSPTTPKSTPLSQNYALT 599
 DB 543 KNGYQSTMASYCAKWKGVDPDSNMAEPKVPPTPINLLGSARSPTTPKSTPLSQNYALT 601
 QY 600 LASDLE-DLALPWSPTNPVAGTAETQNTGEGAGSKACODGQL---SPTWSEIEDLRAAC 655
 DB 602 IQEDLDLALPWSPTNPVAGTAETQNTGEGAGSKACODGQL---SPTWSEIEDLRAAC 654
 QY 656 FG 657
 DB 655 FG 656

RESULT 7
 AAY58164
 ID AAY58164 standard; protein; 550 AA.

XX AC AAY58164;

XX DT 07-MAR-2000 (first entry)

XX DE Adeno associated virus AAV5 Rep68 protein.

XX KW Adeno associated virus; AAV5; AAV2; inverted terminal repeat; ITR; promoter; Rep protein; capsid protein; regulation; transfection; replication; chromosomal integration; tissue tropism; cellular receptor;

KW gene therapy; neutralising antibody; erythroid progenitor cell;
KW transduction; cancer; genetic disease; Rep68.
OS Adeno-associated virus 5.
XX WO9961601-A2.
XX PD 02-DEC-1999.
XX PF 28-MAY-1999; 99WO-US0111958.
XX PF 28-MAY-1998; 98US-0087029P.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PI Chiorini JA, Kotin RM;
XX WPI; 2000-062707/05.
DR N-PSDB; AA249216.
XX Adeno-associated virus 5 based vectors and particles, useful for gene
XX therapy.
XX Claim 17; Page 87-88; 91pp; English.
XX This sequence represents the Rep68 protein of adeno associated virus type
CC 5 (AAV5). The invention relates to vectors comprising a pair of AAVS
CC inverted terminal repeats (ITRs) with a promoter between the ITRs. The
CC vector may comprise the viral genome, or subregions thereof, including
CC sequences encoding Rep proteins and capsid proteins, and is encapsidated
CC in an AAV5 particle. The non-structural Rep proteins Rep40 (AAV58613),
CC Rep52 (AAV58168), Rep68 (AAV58164) and Rep78 (AAV58159) are involved in
CC regulation of replication and transcription, in addition to the
CC production of progeny genomes. Rep68 and Rep78 are also associated with
CC the stable integration of the viral genome into human chromosomes. The
CC three types of capsid protein VP1 (AAV58160), VP2 (AAV58161) and VP3
CC (AAV58162) assemble to form an icosahedral capsid, and differ from each
CC other by the use of alternative splicing and an unusual translation
CC initiation codon (in VP2). AAV5 capsid protein is distinct from AAV2
CC capsid protein and exhibits different tissue tropism. AAV2 and AAV5 are
CC likely to utilise distinct cellular receptors and are serologically
CC distinct. In a gene therapy application, therefore, AAV5 would allow for
CC transduction of a patient who already possess neutralising antibodies
CC either as a result of natural immunological defence or from prior
CC exposure to AAV2 vectors. The vectors may be useful for transducing
CC erythroid progenitor cells or cells lacking heparin sulphate
CC proteoglycans, which is very inefficient with AAV2-based vectors. The
CC vectors may also be useful for transducing cells with a nucleic acid of
CC interest in order to produce cell lines that could be used to screen for
CC agents that interact with the gene product of the nucleic acid of
CC interest. In addition to transduction of other cell types, transduction
CC of erythroid cells would be useful or the treatment of cancer and genetic
CC diseases which can be corrected by bone marrow transplants using matched
CC donors
XX
XX Sequence 550 AA;
SQ Query Match 12.5%; Score 455.5; DB 3; Length 550;
Beat Local Similarity 27.8%; Pred. NO. 1e-34;
Matches 162; Conservative 78; Mismatches 217; Indels 125; Gaps 23;
QY 67 WQSDMEWETTVDEMTEKQVIFEDSLVKKCLFEVLNTKNIFFPGDVNMFVQHEWGKQGW 126
DB 35 WELPESDLNLTVE--QPQLTVADRIRRVLYE---WNKFSKQESKFFVQFKGSEV-F 88
QY 127 HCHVLIGKDF-SQAQGWRRQLNVYWSRWLVLTACNVLTAPRIKLEI-ADENWVT 184
DB 89 HLTLVETSGISSMVLGRY-----VSQIRAOVLV---KVPFGQIEPQINDMVA 132
QY 185 LLYTKHKQTKDYTKVCLFGNMIAYVELTKKISTSPDRDGGYFLSSDSGNK-TNF---- 239
DB 133 I-----TKVKGKGNKVDGSIPIAYLPK-----VQPELOWAWTNLDEYK 173

QY 240 -----LKEGERHLVSKLYTDDMRPETVETVTTTAQETKGR-----IOTKKEVAIKTT 287
DB 174 LAALNLEERKELVAQFLAESSQR-----SQEAASQREFSADPVIKSKTSQKYMAL 223
QY 288 LKELVHKRVTSPEDEMMQPDSDYIEMMAQPGGNNLLKNTLEICTLTARTKTAFDLILEK 347
DB 224 VNLVEHGITTSEKOWIQENQESYLSFNSTGNSRSQIKAAALDNATKIMSLTKSAVDY---- 279
QY 348 AETSKLINFSLPD-----TRTCRIFAFHGN--YV-KVCHAICCVLNRQGGKRNVTFLFHP 400
DB 280 -----LVGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGC---QRSFNKENTVWLYGP 331
QY 401 ASTGKSIIAQAIAQAVGNVCYNAANVPNDCTNKNLIWVEAGNFGQOVNOFKAICS 460
DB 332 ATTGKTNIAEAIAHTVPYPYGCNVNTNENFPDNCVDRKMLIWWEBGKTNKVYESAKAILG 391
QY 461 GQTRIDOKGKSGQIEPTPVIMTNTNENITVVRIGCEERPEHTOPIRDRMLNIHLTHLPL 520
DB 392 GSKVRVDQCKSSQVIDSTPVIIVTSNTNMCVVVDGNSTTFEHQQPLEDRMFKFELTKRLP 451
QY 521 GDFGLVDKNEMPMICAWLVKNGYOSTMASYCAKWKVKPDWSE-----NNAEPKVP 570
DB 452 PDFGKITQEVKDFPAW-----AKYQVPTHEFKVPRELACTGKAESLKL 497
QY 571 TPI-----NLGARSAPFT--TPKSTPLSONYA-LTPL 600
DB 498 RPLGVDVNTSVKLEKRLARLSFVPETPRSSDVTVDPAPLRPL 539
RESULT 8
AAU11409
ID AAU11409 standard; protein; 550 AA.
XX AAU11409;
XX AC AAU11409;
XX DT 26-FEB-2002 (first entry)
XX DE Adeno-associated virus 5 (AAV5), Rep68 protein.
XX KW Adeno-associated virus 5; AAV5; Rep68; nontropic; neuroprotective;
KW cytosstatic; gene therapy; Parkinson's disease; Alzheimer's disease;
KW demyelination disease; metabolic disorder; musculoskeletal disease;
KW cardiovascular disease; cancer; autoimmune disorder; genetic disease;
KW cystic fibrosis; pseudohypoparathyroidism; motile cilia syndrome;
KW bronchitis; pneumonia; emphysema; pulmonary oedema;
KW central nervous system; replication; transcription.
XX Adeno-associated virus 5.
OS WO200170276-A2.
XX PN 27-SEP-2001.
XX PD 22-MAR-2001; 2001WO-US009123.
XX PF 22-MAR-2000; 2000US-00533427.
XX PR (IOWA) UNIV IOWA RES FOUND.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PI Chiorini JA, Kotin RM, Davidson B, Zabner J;
XX WPI; 2002-055104/07.
DR N-PSDB; AAS17712.
XX Delivering nucleic acid into cell for treating Parkinson's disease, by the
PT administering to cell an adeno-associated virus 5 particle comprising the
PT nucleic acid inserted between a pair of AAV inverted terminal repeats.
XX Disclosure; Page 126-127; 130pp; English.
XX The invention describes a novel method of delivering a nucleic acid into
CC a cell in a subject, comprising administering to the cell an adeno-

CC associated virus 5 (AAV5) particle. AAV5 is a small non-pathogenic virus which relies on a helper virus for replication, in the absence of which the AAV5 genome is integrated into a host chromosome in a locus specific manner. The method provides a way to deliver a nucleic acid to a specific regions, tissues and cell types of the central nervous system comprising inserting the nucleic acid between a pair of AAV inverted terminal repeats or delivering an AAV5 particle containing a vector comprising the nucleic acid. The method is useful for treating brain disorders such as demyelination disease, Alzheimer's disease and Parkinson's disease, and metabolic disorders such as musculoskeletal diseases, cardiovascular disease, cancer and autoimmune disorders, for treating genetic diseases such as cystic fibrosis, alpha-1-antitrypsin, pseudohypoadosteronism, imotile cilia syndrome, and for treating bronchitis, pneumonia, emphysema, and cardiogenic and non-cardiogenic pulmonary oedema. AAV5 is useful for delivering gene that may have a systematic effect like anti-hypertension drugs, insulin, coagulation factors, antibiotics, growth factors and hormones. This is the amino acid sequence of the adeno-associated virus 5 (AAV5) Rep68 protein, one of 4 Rep proteins that regulate replication and transcription of the AAV5 genome, described in the method of the invention

XX
CC
SQ Sequence 550 AA;

Query Match 12.5%; Score 455.5; DB 5; Length 550;
Best Local Similarity 27.8%; Pred. No. 1e-34;
Matches 162; Conservative 78; Mismatches 217; Indels 125; Gaps 23;

QY 67 WQSDMEWETTVDKTKQVFIFDLSLVKKCLFEVLNTKNIFFGDVNVNFVQHEWKGDOGW 126
DB 35 WELPPESDLNLTVE--QPOLTVADRIRRVFLYE---WNFKSQESKFFVQFKGSEY-F 88

QY 127 HCHVLIGGKDF-SQAQCKWRRQLNVYMSRWLTACNVQLTPAERIKLREI-AEDNEWVT 184
DB 89 HLHVLTVETSGISSMVLGRY-----VSQIRALV---KVVFQGIETPINDWA 132

QY 185 LLYYKHQTKDYTKCVLFGNMIAYELTKKLTSTSPRGGYFLSDSGWK-TNP----- 239
DB 133 I-----TKVKKGGANKVVDGYPAYLLPK-----VQPELOWMTNLDYK 173

QY 240 -----LKEGERHLVSKLYTDMRPETVTTTAQTKRGR-----IOTKKEVAIKTT 287
DB 174 LAALNLEERKLVQAFLAESQR-----SQEAASQREPSADPVIKSTSQKYMAL 223

QY 288 LKELVHKRVTSPEWMMQPDSEYEMMAQPGENLLKNTLEICTLTARTKTAPDLILEK 347
DB 224 VNWLVHGHTSEKQWIQENQESLTSFNSTGNSRSQIKALDNTATKIMSLTKSAVDY---- 279

QY 348 AETSCLTNFSLPD-----TRTCRIFAFGWN--YV-KVCHAICCVLNRQGGKRVTVLPHGP 400
DB 280 -----LVGSSVPEDISKNRITWQIFEMNGYDPAYAGSILYGCW---QKSFNKRNTVWLYGP 331

QY 401 ASTGKSIIAQIAQAVGNVGCYNAANVFPNDCTKNLWVEAGNFGQOVNQFKAICS 460
DB 332 ATTGKTNIAALAHUTVPFYGCVWNTNENFPNDCVDKRLIWEEGKNTKNVESAAILG 391

QY 461 GQTRIDQKQKSGKQIBPTFVIMTNTENITVVRIGCREPERHTQPIRDRMLNHLTHLTP 520
DB 392 GSKVRVDQKSSQVIDSTPVIIVTSNTNMCVVDGNSSTTFEHQHPLEDRMFKELTKRLP 451

QY 521 GDFGLVDKNWPMICAWLVKNGYQSTWASCAKKGKVPDWSB-----NWAEPKVP 570
DB 452 PDFGKITQKQVDFFAW-----AKVNVQVPVTHEFKVPRELATGKAEKSLK 497

QY 571 TPI-----NLTGARSAPT--TPKSTPLSONYA-LTPL 600
DB 498 RPLGDVNTNTYSKLEKRLAUSFPETPRSSDVTVDPAFLRPL 539

RESULT 9
AAV58159
ID AAV58159 standard; protein; 610 AA.
XX
AC AAV58159;

XX 07-MAR-2000 (first entry)
DT Adeno associated virus AAV5 Rep78 protein.
DE
XX
XX Adeno associated virus; AAV5; AAV2; inverted terminal repeat; ITR;
KW promoter; Rep protein; capsid protein; regulation; transduction;
KW replication; chromosomal integration; tissue tropism; cellular receptor;
KW gene therapy; neutralising antibody; erythroid progenitor cell;
KW transduction; cancer; genetic disease; Rep78.
XX
OS Adeno-associated virus 5.
XX
XX WO9961601-A2.
XX
XX 02-DEC-1999.
XX
XX 28-MAY-1999; 99WO-US011958.
XX
XX 28-MAY-1998; 98US-0087029P.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Chiorini JA, Kotin RM;
XX WPI; 2000-062707/05.
DR N-PSDB; AAZ49211.
DR
XX Adeno-associated virus 5 based vectors and particles, useful for gene therapy.
PT
XX
XX Claim 15; Fig 6; 91pp; English.

This sequence represents the Rep78 protein of adeno associated virus type 5 (AAV5). The invention relates to vectors comprising a pair of AAV5 inverted terminal repeats (ITRs) with a promoter between the ITRs. The vector may comprise the viral genome, or subregions thereof, including sequences encoding Rep proteins and capsid proteins, and is encapsidated in an AAV5 particle. The non-structural Rep proteins Rep40 (AAV58613), Rep52 (AAV59168), Rep68 (AAV58164) and Rep78 (AAV58159) are involved in regulation of replication and transduction, in addition to the production of progeny genomes. Rep88 and Rep78 are also associated with the stable integration of the viral genome into human chromosomes. The three types of capsid protein VP1 (AAV58160), VP2 (AAV58161) and VP3 (AAV58162) assemble to form an icosahedral capsid, and differ from each other by the use of alternative splicing and an unusual translation initiation codon (in VP2). AAV5 capsid protein is distinct from AAV2 capsid protein and exhibits different tissue tropism. AAV2 and AAV5 are likely to utilise distinct cellular receptors and are serologically distinct. In a gene therapy application, therefore, AAV5 would allow for transduction of a patient who already possess neutralising antibodies either as a result of natural immunological defence or from prior exposure to AAV2 vectors. The vectors may be useful for transducing erythroid progenitor cells or cells lacking heparin sulphate proteoglycans, which is very inefficient with AAV2-based vectors. The vectors may also be useful for transducing cells with a nucleic acid of interest in order to produce cell lines that could be used to screen for agents that interact with the gene product of the nucleic acid of interest. In addition to transduction of other cell types, transduction of erythroid cells would be useful or the treatment of cancer and genetic diseases which can be corrected by bone marrow transplants using matched donors

XX SQ Sequence 610 AA;

Query Match 12.5%; Score 455.5; DB 3; Length 610;
Best Local Similarity 27.8%; Pred. No. 1.2e-34;
Matches 162; Conservative 78; Mismatches 217; Indels 125; Gaps 23;

QY 67 WQSDMEWETTVDKTKQVFIFDLSLVKKCLFEVLNTKNIFFGDVNVNFVQHEWKGDOGW 126
DB 35 WELPPESDLNLTVE--QPOLTVADRIRRVFLYE---WNFKSQESKFFVQFKGSEY-F 88

QY 127 HCHVLIGKDF -SQAQGWRRQLNVYNSRWLVTACNVQLTPAERIKLREI-AEDNEWTT 184
DB 89 HLHVLVETSGISMWLGRY-----VSQIRAOVLV---KVFGIEPQINDWA 132
QY 185 LLTYKHQKDYTKCVLFGNMIAYFLTKKISTSPRGGYFLSSDGMK-TNF----- 239
DB 133 I-----TKVKGKANKVDSGYIPAYLLPK-----VQPELOWAWTNLDEYK 173
QY 240 -----LKEGERHLVSKLYTDDMRPEVTVTVAQTKRGR-----IOTKKEVAIKTT 287
DB 174 LAALNLEERKRLVAQFLAESSQR-----SQAASQREFSADPVIKSTKSQKYMAL 223
QY 288 LKELVHKRVTSPEWMMQPDYSIEMMAQPGENLKNLTLEICTLTLARTKTAFLDILEK 347
DB 224 VNWLVHEGITSEKOWIQENQESYLSFNSTGNSRSQIKALDNATKIMSLTKSAVDY---- 279
QY 348 AETSCLTNFSLPD-----TRTCRIFAFHGMN--YV-KVCHAI CCVLNRQGGKNTVLFHGP 400
DB 280 -----LVGSSVPEDISKNRWQIFEMNGYDPAYAGSILYGC-----QRSFNKNTVWLYGP 331
QY 401 ASTGKSIIAQIAQAVGNCVNAANVPFNDCNTKNLIWVEEAGNFQGVNQFKAICS 460
DB 332 ATTGKTNIABIAIAHTVPFYGCNVNTNENFPFNDVCDKMLIWEEGKNTKVVESAKAILG 391
QY 461 GQTRIDQKGGKSKQIEPTVIMTNTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLP 520
DB 392 GSKVRVDQCKSSVQIDSTPVI VTSNTNMCVVVDGNSSTTFEHOQPLEDRMFELTKRLP 451
QY 521 GDFGLVDKNEWPMICAWLVKNGYQSTWASACAKGKVPDWSE-----NWAEPKVP 570
DB 452 PDFGKITKQEVKDFFAW-----AKVNVQVPTHEFKVPRELAGTGAEKSLK 497
QY 571 TPI-----NLLGARSFPFT--TPKSTPLSONYA-LTPL 600
DB 498 RPLGDVNTSYKSLEKEARLSFVPEPRSSDVTVDPAPLRPL 539

RESULT 10

AAI97720 standard; protein; 610 AA.
ID AAY97720 standard; protein; 610 AA.
AC AAY97720;
XX 19-JUN-2001 (first entry)
DT Rep protein sequence.
XX Fusion nucleic acid library; Rep protein; tumour cell; apoptosis;
KW nucleic acid modification enzyme; cell death; decreased cell growth;
KW protein-protein interaction detection; cell division; cancer therapy;
KW protein drug discovery; pharmacogenetics.
XX Adeno associated virus 5.
OS WO200114539-A2.
XX 01-MAR-2001.
XX 18-AUG-2000; 2000WO-US022906.
XX 20-AUG-1999; 99US-0150004P.
XX 02-JUN-2000; 2000US-0209130P.
XX (UYJO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX Li M;
XX WPI; 2001-218443/22.
XX N-PSDB; AAA91310.

XX New library of fusion nucleic acids each encoding a Rep protein
PT recognized by a nucleic acid modification enzyme and a candidate protein,
PT useful for detecting protein-protein interactions, protein drug discovery

or pharmacogenetics.

Disclosure; Fig 21; 106pp; English.

This sequence is the adeno associated virus 5 Rep protein. The invention relates to a library of fusion nucleic acids, each encoding a Rep protein, a candidate protein, a presentation structure, a targeting sequence or a label. The Rep protein is a nucleic acid modification enzyme. The random or directed libraries (including the cDNA libraries) can be introduced into any tumour cell, and peptides identified which by themselves induce apoptosis, cell death, loss of cell division or decreased cell growth. The methods and compositions may also be used to detect protein-protein interactions, protein drug discovery, particularly for protein drugs that interact with targets on cell surfaces, to discover DNA or nucleic acid binding proteins, using nucleic acids as targets, to screen for nucleic acid modification enzymes with decreased toxicity for the host cells, to identify or generate Rep proteins with expression vectors and in pharmacogenetic studies. The method is useful in cancer therapy and in killing tumour cells. The methods can be combined with other cancer therapeutics (drugs or radiation) to sensitize cells and thus induce rapid and specific apoptosis, cell death, loss of cell division or decreased cell growth after exposure to a secondary agent

XX SQ Sequence 610 AA;

Query Match 12.5%; Score 455.5; DB 4; Length 610;
Best Local Similarity 27.8%; Pred. No. 1.2e-34;
Matches 162; Conservative 78; Mismatches 217; Indels 125; Gaps 23;

QY 67 WQOSDMWEITVDENTKKQVFPDLSVKCLFEVLNTKNIFPGDVNMFVQHEWKGDOGW 126
DB 35 WELPPESDLNLTVE--QPULTVADRIRRVFLYE---WNKFSKQBSKEFVFEKGEY-F 88
QY 127 HCHVLIGKDF -SQAQGWRRQLNVYNSRWLVTACNVQLTPAERIKLREI-AEDNEWTT 184
DB 89 HLHVLVETSGISMWLGRY-----VSQIRAOVLV---KVFGIEPQINDWA 132
QY 185 LLTYKHQKDYTKCVLFGNMIAYFLTKKISTSPRGGYFLSSDGMK-TNF----- 239
DB 133 I-----TKVKGKANKVDSGYIPAYLLPK-----VQPELOWAWTNLDEYK 173
QY 240 -----LKEGERHLVSKLYTDDMRPEVTVTVAQTKRGR-----IOTKKEVAIKTT 287
DB 174 LAALNLEERKRLVAQFLAESSQR-----SQAASQREFSADPVIKSTKSQKYMAL 223
QY 288 LKELVHKRVTSPEWMMQPDYSIEMMAQPGENLKNLTLEICTLTLARTKTAFLDILEK 347
DB 224 VNWLVHEGITSEKOWIQENQESYLSFNSTGNSRSQIKALDNATKIMSLTKSAVDY---- 279
QY 348 AETSCLTNFSLPD-----TRTCRIFAFHGMN--YV-KVCHAI CCVLNRQGGKNTVLFHGP 400
DB 280 -----LVGSSVPEDISKNRWQIFEMNGYDPAYAGSILYGC-----QRSFNKNTVWLYGP 331
QY 401 ASTGKSIIAQIAQAVGNCVNAANVPFNDCNTKNLIWVEEAGNFQGVNQFKAICS 460
DB 332 ATTGKTNIABIAIAHTVPFYGCNVNTNENFPFNDVCDKMLIWEEGKNTKVVESAKAILG 391
QY 461 GQTRIDQKGGKSKQIEPTVIMTNTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLP 520
DB 392 GSKVRVDQCKSSVQIDSTPVI VTSNTNMCVVVDGNSSTTFEHOQPLEDRMFELTKRLP 451
QY 521 GDFGLVDKNEWPMICAWLVKNGYQSTWASACAKGKVPDWSE-----NWAEPKVP 570
DB 452 PDFGKITKQEVKDFFAW-----AKVNVQVPTHEFKVPRELAGTGAEKSLK 497
QY 571 TPI-----NLLGARSFPFT--TPKSTPLSONYA-LTPL 600
DB 498 RPLGDVNTSYKSLEKEARLSFVPEPRSSDVTVDPAPLRPL 539

RESULT 11

a fusion of nucleic acids encoding NAM enzyme and the compound.

Disclosure; Fig 21; 96pp; English.

The present invention relates to genetic libraries of nucleic acid/protein (NAP) conjugates comprising a fusion polypeptide (with a nucleic acid modification (NAM) enzyme (E) and candidate compound), an expression vector (with a fusion of nucleic acids encoding the enzyme and candidate protein respectively), an enzyme attachment sequence (EAS; RNA sequence), where the candidate compound and candidate protein are different and EAS and the enzyme are covalently linked. The NAP conjugates are useful in screens to assay binding to target molecules and/or to screen candidate agents for the ability to modulate the activity of the target molecule. They are useful in cancer therapy. Sequences of the invention are also useful to detect protein-protein interaction, in drug discovery, to discover DNA or nucleic acid binding proteins, using nucleic acids as the targets and to screen for NAM enzymes with reduced toxicity. NAP conjugates are also useful in pharmacogenomic studies, for screening bioactive cells (specifically Rep proteins with reduced toxicity). NAP conjugates are also useful in screening proteins causing phenotypic changes such as overproduction or inhibition of protein expression, or proteins that alter attachment, infectivity, etc. of the virus. Sequences of the invention are also used in gene therapy. The present sequence is adeno-associated virus (AAV) 5 Rep. (Updated on 29-AUG-2003 to standardise OS field)

Sequence 610 AA;

Query Match 12.5%; Score 455.5; DB 5; Length 610;

Best Local Similarity 27.8%; Pred. No. 1.2e-34;

Matches 162; Conservative 78; Mismatches 217; Indels 125; Gaps 23;

67 WDQSEDEWETTTVDKTKQVIFDLSLVKKCLPEVLNTKNIIPGDNVNVFVQHEWKGQDQW 126
35 WELPPESDLNLTVE--QPOLTVADRIRRVFLVE---WNKFSKQESKFFVQFEKSEY-F 88
127 HCHVLIGGKDF-SQAQGWRRQLNVYWSRWLTACNVOLTPAERIKLREI-AEDNEWVT 184
89 HLHVLTVETSGISSMVLGRY-----VSQIRALV---KVVFQGLEPQINDWA 132
185 LLTYKHKQTKDYTKCVLFGNMIAFYFLTKKISTSPRGGVPLSSDSGWK-TNP---- 239
133 I-----TKVKGKGVNDVSGYIPAYLLPK-----VQPELQWANTNLDEYK 173
240 -----LKEGERHLVSKLVTDDMRPETVETTTTAQETKGR-----IQTKKEVALKT 287
174 LAALNLEERKLVAQFLAESSOR-----SQEAASQREFSADPVIKSKTSQKTMAL 223
288 LKELVHKVETSPEDMMWQPDYSIEMMAQPGENLLKNTLICTLTILARTKTAFLILEK 347
224 VNWLVHEGHTSEKQWIOENQESYLSFNSTGNSRSGIKAAALDNATKIMSLYKSAVDY---- 279
348 AETSCLTNFSLPD-----TRTCIRAFHGNW--YV-KVCHAJCCVLNFGGGRNTVLPHGP 400
280 -----LVGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGC---QRSFNKNTVWLYGP 331
401 ASTGKSIITAAQIAQVGNVGNVNAANVFPNDCTNKULIWVEEAGNFQGVQVQKALCS 460
332 ATTKGNTIAEAIAHTVPYGYCVNTNENFPFNDVCDVKMLIWVEGKMTKNVVEAKALIG 391
461 GQTRIDOKGSGKQIEPTTPVIMTNNITVVRIGCEERPEHTQPIRDRMLNIHLTHLP 520
392 GSKVRDOKCKSGVQIDSTPVLVTSNTNMVVDGNSITFEHQPLEDRMFKFELTKRLP 451
521 GDFGLVDKNEWPMICAWLVKNGYQSTMASYCAKWKGVDFNDE-----NNAEPKVP 570
452 PDFGKITTKQEVKDFPAM-----AKVQVFTVTHEFKVPRELACTKGAESLK 497
571 TPI-----NILGARSPT--TPKSTPLSQNVA-LTLP 600
498 RPLGDVNTNTSYKLEKRLARLSFVPETPRSSDVTDVDFLRPL 539

RESULT 13

RAE28644

ID AAE28644 standard; protein; 610 AA.

XX AC AAE28644;

XX 29-AUG-2003 (revised)

DT 27-DEC-2002 (first entry)

XX Adeno-associated virus 5 Rep protein.

XX Nucleic acid modification enzyme; NAM; enzyme attachment sequence; EAS;
KW protein design automation; PDA; cancer; protein-protein interaction;
KW infection; gene therapy; Rep protein.
XX Adeno associated virus; 5.
XX WO200268453-A2.
XX 06-SEP-2002.
XX 19-FEB-2002; 2002WO-US004853.
XX 22-FEB-2001; 2001US-007926229.
XX (XENC-) XENCOR INC.
XX Li M, Dahiyat BI;
XX WPI; 2002-691653/74.
XX N-PSDB; AAD46138.

XX Generating a library of fusion nucleic acids for treating cancer or

XX infection, or detecting protein-protein interaction, comprises providing

XX computationally-derived library of candidate protein sequences and

XX expression vectors.

XX Disclosure; Page 180-182; 246pp; English.

XX The present invention relates to a novel method of generating a library
XX of fusion nucleic acids. The method involves providing a computationally-
XX derived library of candidate protein sequences and creating a library of
XX expression vectors containing a fusion nucleic acid having a sequence
XX encoding a nucleic acid modification (NAM) enzyme and a sequence encoding
XX a candidate protein sequence from the library and an enzyme attachment
XX sequence (EAS) that is recognised by the NAM enzyme. The invention also
XX relates to the use of a variety of computation methods including protein
XX design automation (PDA). The method is useful in generating and screening
XX fusion nucleic acids that may be used in treating cancer or infections,
XX in detecting protein-protein interactions, discovery of DNA or nucleic
XX acid binding proteins, protein drug discovery, screening for NAM enzymes
XX with decreased toxicity to the host cells and NAM enzyme/EAS pairs with
XX increased affinity or in pharmacogenetic studies. The invention is also
XX used in gene therapy. The present sequence is Adeno-associated virus 5
XX Rep protein. This sequence is used to illustrate the method of the
XX invention. (Updated on 29-AUG-2003 to standardise OS field)

Sequence 610 AA;

Query Match 12.5%; Score 455.5; DB 5; Length 610;

Best Local Similarity 27.8%; Pred. No. 1.2e-34;

Matches 162; Conservative 78; Mismatches 217; Indels 125; Gaps 23;

Qy 67 WDQSEDEWETTTVDKTKQVIFDLSLVKKCLPEVLNTKNIIPGDNVNVFVQHEWKGQDQW 126

Db 35 WELPPESDLNLTVE--QPOLTVADRIRRVFLVE---WNKFSKQESKFFVQFEKSEY-F 88

Qy 127 HCHVLIGGKDF-SQAQGWRRQLNVYWSRWLTACNVOLTPAERIKLREI-AEDNEWVT 184

Db 89 HLHVLTVETSGISSMVLGRY-----VSQIRALV---KVVFQGLEPQINDWA 132

Qy 185 LLTYKHKQTKDYTKCVLFGNMIAFYFLTKKISTSPRGGVPLSSDSGWK-TNP---- 239

Db 133 I-----TKVKGANKVDSGYIPAYLLPK-----VQPELOWANTNLDYK 173
Qy 240 -----LKEGERHLVSKLYTDDMRPETVETVTTTAQETKGR-----IQTKKEVAIKTT 287
Db 174 LAALNLEERKRLVAQFLAESSQR-----SQEAASQREFSADPVIKSKTSQKYMAL 223
Qy 288 LKELVHKRVTSPEWMMQPDSEIENMAQPGGKLLKNTLEICTLTLARTKTAFLDILEK 347
Db 224 VNWLVHGGITSEKQIQENQESYLSFNSTGNSRSQIKAAALDNATKIMSLTKSAVDY---- 279
Qy 348 AETSCLTNFSLPD-----TRTCRIFAFHGMN--YV-KVCHAI CCVLRQGGKRVTLFHPG 400
Db 280 -----LVGSSVPEDISKRIWQIFENMGYDPAYAGSILYGC-----QRSFNKRVTVWLYGP 331
Qy 401 ASTGKSIIAQAIAQAVGNCYNAANVFPFNDCTNKNLIWVEEAGNFGQOVNQFKAICS 460
Db 332 ATTGKTNI AEIAHTVPFYGCNVNWTNENFPFNDCTNKNLIWVEEAGNFGQOVNQFKAICS 391
Qy 461 GQIRIDQKSGSKQIEPTVIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPL 520
Db 392 GSKVRVDQKSSQVQIDSTPEVITVNTNMCVVVDGNSSTTFEHOQPLEDRMFKFELTKRLP 451
Qy 521 GDFGLVDKNEWPMICAWLVKNGYQSTWASVCAKWKGVDPWSE-----NWABPKVP 570
Db 452 PDFGKITQEVKDFFAW-----AKVQVPTVTHEFKVPRELAGTKGAESKLK 497
Qy 571 TPI-----NLLGSARSPT--TPKSTPLSQNYA-LTPL 600
Db 498 RPLGVDVNTSYKSLKRLARLSFVPEPRSSDVTVDPAPLRPL 539

RESULT 14
AAE26940
ID AAE26940 standard; protein; 610 AA.
XX AC AAE26940;
XX DT 13-DEC-2002 (first entry)
DE Adeno associated virus 5 Rep protein.
XX Prokaryotic library; candidate protein; nucleic acid modification; NAM;
KW enzyme attachment sequence; EAS; clinical pharmacology; chemical sensor;
KW enzymology; cosmetic research; toxic; environmental safety assessment;
KW nutrient biology; Rep protein.
XX Adeno associated virus.
XX WO200266653-A2.
PN 14-DEC-2001; 2001WO-US049058.
XX 29-AUG-2002.
PF 14-DEC-2001; 2001WO-US049058.
XX 14-DEC-2000; 2000US-0256163P.
PR (XENC-) XENCOR INC.
XX Li M, Liu Y;
PI WPI; 2002-667068/71.
DR N-PSDB; AAD44600.
XX New library of prokaryotic pET-24a expression vectors, host cells or
PT nucleic acid/protein conjugates, useful for screening candidate proteins
PT and their nucleic acids or modification enzymes for pharmacogenetic
PT analysis.
XX Disclosure; Fig 21; 127pp; English.
XX The invention relates to methods and compositions for the construction of
CC prokaryotic libraries expressing candidate proteins and the use of these

CC libraries to identify candidate proteins and the nucleic acids encoding
CC them. The invention provides a library of prokaryotic pET-24a vectors
CC comprising a fusion nucleic acid consisting of a nucleic acid encoding a
CC nucleic acid modification (NAM) enzyme or a candidate protein, or a
CC nucleic acid having a T7 promoter operably linked to the NAM enzyme or
CC the candidate protein, and an enzyme attachment sequence (EAS) recognised
CC by the NAM enzyme. The library is used for identifying candidate proteins
CC and nucleic acids encoding these proteins, in screening for NAM enzymes
CC with decreased toxicity for the host cells, or in identifying novel or
CC improved EASs, which may be used for understanding cellular processes or
CC any subsequent therapeutic or toxic activities. The nucleic acid/protein
CC (NAP) conjugates are useful in diagnostic assays and in research
CC including clinical pharmacology, functional genomics, pharmacogenomics,
CC agricultural chemicals, environmental safety assessment, chemical sensor,
CC nutrient biology, cosmetic research or enzymology. These may also be used
CC in vitro screening techniques and in assays with target molecules. The
CC present sequence is Adeno associated virus 5 Rep protein used in the
CC invention
XX
SQ Sequence 610 AA;
Query Match 12.5%; Score 455.5; DB 5; Length 610;
Best Local Similarity 27.8%; Pred. No. 1.2e-34;
Matches 162; Conservative 78; Mismatches 217; Indels 125; Gaps 23;
Qy 67 WDOSEDMETTVDEMTKKQVFI FDSLIVKCLPEVLNTRKNIFPGDVNVFVQHWGKDGQW 126
Db 35 WELPPESDLNLTIVE--QPQLTVADRIRRVFLYE--WNKFSQESKFFVQFEKSEY-F 88
Qy 127 HCHVLIGGKDF-SQAQGWRRQLNVYWSRWLTACNVOLTPAERIKLREI-AEDNEWVT 184
Db 89 HLHTLVETGSSWMLGRY-----VSQIRAOQLV---KVVFQIEFQINDMVA 132
Qy 185 LLTYKHQTKDYTKCVLFQGNMIAYYFLTKKISTSPRDGGYFLSDSGWK-TNF---- 239
Db 133 I-----TKVKGANKVDSGYIPAYLLPK-----VQPELOWANTNLDYK 173
Qy 240 -----LKEGERHLVSKLYTDDMRPETVETVTTTAQETKGR-----IQTKKEVAIKTT 287
Db 174 LAALNLEERKRLVAQFLAESSQR-----SQEAASQREFSADPVIKSKTSQKYMAL 223
Qy 288 LKELVHKRVTSPEWMMQPDSEIENMAQPGGKLLKNTLEICTLTLARTKTAFLDILEK 347
Db 224 VNWLVHGGITSEKQIQENQESYLSFNSTGNSRSQIKAAALDNATKIMSLTKSAVDY---- 279
Qy 348 AETSCLTNFSLPD-----TRTCRIFAFHGMN--YV-KVCHAI CCVLRQGGKRVTLFHPG 400
Db 280 -----LVGSSVPEDISKRIWQIFENMGYDPAYAGSILYGC-----QRSFNKRVTVWLYGP 331
Qy 401 ASTGKSIIAQAIAQAVGNCYNAANVFPFNDCTNKNLIWVEEAGNFGQOVNQFKAICS 460
Db 332 ATTGKTNI AEIAHTVPFYGCNVNWTNENFPFNDCTNKNLIWVEEAGNFGQOVNQFKAICS 391
Qy 461 GQIRIDQKSGSKQIEPTVIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPL 520
Db 392 GSKVRVDQKSSQVQIDSTPEVITVNTNMCVVVDGNSSTTFEHOQPLEDRMFKFELTKRLP 451
Qy 521 GDFGLVDKNEWPMICAWLVKNGYQSTWASVCAKWKGVDPWSE-----NWABPKVP 570
Db 452 PDFGKITQEVKDFFAW-----AKVQVPTVTHEFKVPRELAGTKGAESKLK 497
Qy 571 TPI-----NLLGSARSPT--TPKSTPLSQNYA-LTPL 600
Db 498 RPLGVDVNTSYKSLKRLARLSFVPEPRSSDVTVDPAPLRPL 539

RESULT 15
ABU64865
ID ABU64865 standard; protein; 610 AA.
XX AC ABU64865;
XX DT 14-MAY-2003 (first entry)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:27:50 ; Search time 25 Seconds
(without alignments)
1782.628 Million cell updates/sec

Title: US-10-069-056-6

Perfect score: 3638

Sequence: 1 MAGNAYSDEVLGATNWLKKEK.....RACFGAEPLKDFSEPLNLD 672

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/prodata/1/iaa/PCOTUS_COMB.pep.*
- 6: /cgn2_6/prodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	437.5	12.0	536	4 US-09-532-594B-10	Sequence 10, Appl
2	437.5	12.0	623	4 US-09-532-594B-2	Sequence 2, Appl
3	437.5	12.0	623	4 US-09-532-594B-11	Sequence 11, Appl
4	433.5	11.9	546	4 US-09-807-802A-7	Sequence 7, Appl
5	433.5	11.9	623	4 US-09-807-802A-2	Sequence 2, Appl
6	433.5	11.9	623	4 US-09-807-802A-5	Sequence 5, Appl
7	430.5	11.8	312	4 US-09-532-594B-8	Sequence 8, Appl
8	430.5	11.8	399	4 US-09-532-594B-9	Sequence 9, Appl
9	425.5	11.7	399	4 US-09-807-802A-11	Sequence 11, Appl
10	425.5	11.7	399	4 US-09-807-802A-9	Sequence 9, Appl
11	144	4.0	276	3 US-09-171-461-3	Sequence 3, Appl
12	144	4.0	276	3 US-09-970-711-3	Sequence 3, Appl
13	122.5	3.4	647	3 US-09-300-909-17	Sequence 17, Appl
14	113	3.1	589	4 US-09-107-532A-6086	Sequence 6086, Ap
15	112	3.1	564	3 US-09-134-001C-2897	Sequence 2897, Ap
16	103	2.8	1503	4 US-09-677-046A-6	Sequence 6, Appl
17	103	2.8	1509	4 US-09-677-046A-2	Sequence 2, Appl
18	101	2.8	481	4 US-09-252-991A-30005	Sequence 30005, A
19	101	2.8	1170	4 US-09-638-524A-2	Sequence 2, Appl
20	100	2.7	1120	4 US-09-792-024-95	Sequence 95, Appl
21	99.5	2.7	2409	6 5180808-2	Patent No. 5180808
22	99	2.7	919	4 US-09-248-796A-19408	Sequence 19408, A
23	98.5	2.7	556	4 US-09-248-796A-18763	Sequence 18763, A
24	98.5	2.7	3248	1 US-08-353-700-1	Sequence 1, Appl
25	98.5	2.7	3248	5 PCT-US95-16216-1	Sequence 1, Appl
26	98	2.7	605	4 US-09-107-532A-6077	Sequence 6077, Ap
27	98	2.7	715	4 US-09-248-796A-20602	Sequence 20602, A

28 97.5 2.7 657 3 US-09-300-909-16 Sequence 16, Appl
29 97 2.7 291 4 US-09-107-532A-6390 Sequence 6390, Ap
30 97 2.7 471 4 US-09-134-000C-6228 Sequence 6228, Ap
31 97 2.7 1115 2 US-08-568-459A-2 Sequence 2, Appl
32 97 2.7 1115 2 US-08-487-826B-2 Sequence 2, Appl
33 97 2.7 1115 3 US-09-210-288-2 Sequence 2, Appl
34 97 2.7 1115 6 5198347-6 Patent No. 5198347
35 96 2.6 370 4 US-09-583-110-3405 Sequence 3405, Ap
36 95.5 2.6 441 4 US-09-540-236-2983 Sequence 2983, Ap
37 95.5 2.6 957 4 US-09-489-039A-11233 Sequence 11233, A
38 95 2.6 590 4 US-09-134-000C-4755 Sequence 4755, Ap
39 95 2.6 1307 1 US-08-395-246C-2 Sequence 2, Appl
40 94.5 2.6 1060 4 US-09-248-796A-16624 Sequence 16624, A
41 94.5 2.6 2482 1 US-08-328-254-6 Sequence 6, Appl
42 94.5 2.6 3210 4 US-09-538-092-1154 Sequence 1154, Ap
43 94 2.6 607 3 US-08-486-099-116 Sequence 116, App
44 94 2.6 607 3 US-08-484-223B-116 Sequence 116, App
45 94 2.6 607 3 US-08-919-597-116 Sequence 116, App

ALIGNMENTS

RESULT 1

US-09-532-594B-10

; Sequence 10, Application US/09532594B

; Patent No. 6468524

; GENERAL INFORMATION:

; APPLICANT: Chorini, John A.

; APPLICANT: Kotin, Robert M.

; APPLICANT: Saffer, Brian

; APPLICANT: Davidson, Beverly

; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF

; FILE REFERENCE: 14014.0252U2

; CURRENT APPLICATION NUMBER: US/09/532,594B

; CURRENT FILING DATE: 2000-03-22

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 10

; LENGTH: 536

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =

; OTHER INFORMATION: synthetic construct

; NAME/KEY: misc feature

; OTHER INFORMATION: AAV4 Rep protein 68

; US-09-532-594B-10

Query Match 12.0%; Score 437.5; DB 4; Length 536;

Best Local Similarity 27.7%; Pred. No. 1.5e-36;

Matches 157; Conservative 75; Mismatches 232; Indels 103; Gaps 19;

QY 63 AETWQSDMEWETTVDEMTKKQVIFDLSLVKKCLFEVLNTKNIIPGDVNVVQHWGK 122

DB 31 AEKEWELPPDSMDLNLIE--QAPLTVAEKLOREFLVEMRRVSK--APEALFFVQFEKG- 85

QY 123 DQGWCHCHVLIG-----GKDFSOAQGKWMRQLNVYMSRWLTACNVQLTPAERIKL 173

DB 86 DSYFHLHLIVETGVKSMVGVRSIQKEK-----LVTRIYGVPE----- 126

QY 174 RETAEDNEWVTLTYKHQ--TKKDYTKCVLFGNMIAYYPLTKKKISTSPRGGVFLSS 231

DB 127 ---QLPNFNAVTKTRNGAGGNKVVDCC-----YIPNLLPK---TQP----- 163

QY 232 DSGWK-TNF-----LKEGRLHVLVKLYTDDMPETVETVTTTAQETKRG----- 275

DB 164 ELQWANTMDQYISACLNLAERKLVQOHL-----THVSOTQEQNKQNPNSD 212

QY 276 ---IQTKKEVAIKTTIKELVHKVETSPEDMMQPDSPYEMMAQPGGELKNTLEICTL 332

DB 213 APVIRKTSARYMELVGNLVDROITSEKQIQEDQASYSIFNAASNSRSQIKAAALDNASK 272

```
QY 333 TLARTKTAFLDILEKAETSKLNFSLPDRTRCFAPFHGWNVYKVAICCVLNRQGGKR 392
Db 273 INSLTKTAPDYLVGQNPDPEDISS-----NRIYRILEMNGYDPQYAAASVFLGWAQKKGKGR 327
QY 393 NTVLFHGPASTGKSIIAQIAOAVGNGVCYNAANVPFNDCTNKNLIWVEEAGNFGQOV 452
Db 328 NTIWLFGPATTKGTNTAEIAHAHVPYGCNVNTNENFPENDCVDKMWIWEESGKMTAKV 387
QY 453 NQFKATCSGQTRIDOKGKSGQIBPTPVIMTNNENITVVRIGCEBEPHPTQPIRDRMLN 512
Db 388 ESAKAILGGSKVRVDQCKSSAQIDPTPVIVTSNTNMCVIDGNSSTTFEHOQPLQDRMPK 447
QY 513 IHLTHLPGDFGLVDKNWPMICAMLVKNGYQSTWASACAKWG--KVPDWSB-NWAEPKV 569
Db 448 FELTKLEHDFGKVTQKQEVKDFRWDASHVTEVTFYVRKGGARKRPAPNDADISEPKR 507
QY 570 PTPINLGSARSPTTTPKSTPLSQNYA 596
Db 508 ACP-----SVAQPSTSDAEPV--DYA 527

RESULT 2
US-09-532-594B-2
; Sequence 2, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Davidon, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.025202
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; TYPE: PRT
; LENGTH: 623
; ORGANISM: Artificial Sequence.
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 Rep protein (full length)
US-09-532-594B-2

Query Match 12.0%; Score 437.5; DB 4; Length 623;
Best Local Similarity 27.7%; Pred. No. 2e-36;
Matches 157; Conservative 75; Mismatches 232; Indels 103; Gaps 19;

QY 63 AETTWDQSEDMEWETTVDENTKKQVFIDSLVKKCLFEVLNTKNIPFGDVNWFVQHEWGK 122
Db 31 AEKEWELPPDSMDLNLIE--QAPLTVAEKLOREFLVEWRVSK--APEALFFVQPEKG- 85
QY 123 DQGWCHVLIG-----GKDFSOAQGWRRQLNVVWSRWLVATCNVLTQPAERIKL 173
Db 86 DSYFHLHILVETGVKSWVGRVYSQIKEK-----LVTRIYRGVEP----- 126
QY 174 RETAEDNEWVTLTYKHQ--TKDYTKCVLFGNMIAYFLTKKKISTSPRGGYFLSS 231
Db 127 ----QLPNWFAVTKTRNGAGGNKVDDC-----YIPNLLPK---TQP----- 163
QY 232 DSGWK-TNF-----LKEGERHLVSKLYTDDMRPETVETVTTTAQETKGR----- 275
Db 164 ELQAWNTMDQYISACLNLAEKRLVAQHL-----THVSOTQEQNKENQNPNSD 212
QY 276 ---IQTKKEVAIKTKLKVHKTSPEDWMMQPSYIEMMAQPGENLLKNTLICTL 332
Db 213 APVIRSKTSARYMELVGLVDRGITSEKQWIOEDQASYISFNAAASRSQIKAAALONASK 272
QY 333 TLARTKTAFLDILEKAETSKLNFSLPDRTRCFAPFHGWNVYKVAICCVLNRQGGKR 392
Db 273 INSLTKTAPDYLVGQNPDPEDISS-----NRIYRILEMNGYDPQYAAASVFLGWAQKKGKGR 327
```

```
Db 273 INSLTKTAPDYLVGQNPDPEDISS-----NRIYRILEMNGYDPQYAAASVFLGWAQKKGKGR 327
QY 393 NTVLFHGPASTGKSIIAQIAOAVGNGVCYNAANVPFNDCTNKNLIWVEEAGNFGQOV 452
Db 328 NTIWLFGPATTKGTNTAEIAHAHVPYGCNVNTNENFPENDCVDKMWIWEESGKMTAKV 387
QY 453 NQFKATCSGQTRIDOKGKSGQIBPTPVIMTNNENITVVRIGCEBEPHPTQPIRDRMLN 512
Db 388 ESAKAILGGSKVRVDQCKSSAQIDPTPVIVTSNTNMCVIDGNSSTTFEHOQPLQDRMPK 447
QY 513 IHLTHLPGDFGLVDKNWPMICAMLVKNGYQSTWASACAKWG--KVPDWSB-NWAEPKV 569
Db 448 FELTKLEHDFGKVTQKQEVKDFRWDASHVTEVTFYVRKGGARKRPAPNDADISEPKR 507
QY 570 PTPINLGSARSPTTTPKSTPLSQNYA 596
Db 508 ACP-----SVAQPSTSDAEPV--DYA 527

RESULT 3
US-09-532-594B-11
; Sequence 11, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Davidon, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.025202
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
; TYPE: PRT
; LENGTH: 623
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 Rep protein 78
US-09-532-594B-11

Query Match 12.0%; Score 437.5; DB 4; Length 623;
Best Local Similarity 27.7%; Pred. No. 2e-36;
Matches 157; Conservative 75; Mismatches 232; Indels 103; Gaps 19;

QY 63 AETTWDQSEDMEWETTVDENTKKQVFIDSLVKKCLFEVLNTKNIPFGDVNWFVQHEWGK 122
Db 31 AEKEWELPPDSMDLNLIE--QAPLTVAEKLOREFLVEWRVSK--APEALFFVQPEKG- 85
QY 123 DQGWCHVLIG-----GKDFSOAQGWRRQLNVVWSRWLVATCNVLTQPAERIKL 173
Db 86 DSYFHLHILVETGVKSWVGRVYSQIKEK-----LVTRIYRGVEP----- 126
QY 174 RETAEDNEWVTLTYKHQ--TKDYTKCVLFGNMIAYFLTKKKISTSPRGGYFLSS 231
Db 127 ----QLPNWFAVTKTRNGAGGNKVDDC-----YIPNLLPK---TQP----- 163
QY 232 DSGWK-TNF-----LKEGERHLVSKLYTDDMRPETVETVTTTAQETKGR----- 275
Db 164 ELQAWNTMDQYISACLNLAEKRLVAQHL-----THVSOTQEQNKENQNPNSD 212
QY 276 ---IQTKKEVAIKTKLKVHKTSPEDWMMQPSYIEMMAQPGENLLKNTLICTL 332
Db 213 APVIRSKTSARYMELVGLVDRGITSEKQWIOEDQASYISFNAAASRSQIKAAALONASK 272
QY 333 TLARTKTAFLDILEKAETSKLNFSLPDRTRCFAPFHGWNVYKVAICCVLNRQGGKR 392
Db 273 INSLTKTAPDYLVGQNPDPEDISS-----NRIYRILEMNGYDPQYAAASVFLGWAQKKGKGR 327
```

QY 393 NTVLPHGPASTGKSIIAQIAQAVGNVGCYNAANVPFPNDCTNKNLIWVEEAGNFGQV 452
Db 328 NTWLFPGPATGTGNTIAEIAHAHVPYGCNVNTNENFPFNDCTNKNLIWVEEAGNFGQV 387
QY 453 NQFKALCSGQIRIDQKSGKQIEPTPVIMTNNENITVVRIGCEERPEHTQPIRDRMLN 512
Db 388 ESAKAILGSKVRVDQCKSSAQIDPTPLVTSNTMCAVIDGNSITTFEHOQLQDRMPK 447
QY 513 IHLTHLPDGLVLDKNEWPMICAWLVKNGYQSTMASYCAKMG--KVPDMS--NWAEPKV 569
Db 448 FELTKRLEHDFGKVKQEVKDFRWSADHVTETHEFYVRKGGARKRPAPNDADISEPKR 507
QY 570 PTPINLLGSARSPTTPKSTPLSQNYA 596
Db 508 ACP-----SVAQSTSDAEPV--DYA 527
RESULT 4
US-09-807-802A-7
; Sequence 7, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNVN.031USA
; CURRENT FILING DATE: 2002-02-21
; PRIOR FILING DATE: 2002-02-21
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 546
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-7

Query Match 11.9%; Score 433.5; DB 4; Length 546;
Best Local Similarity 27.0%; Pred. No. 4.2e-36;
Matches 157; Conservative 63; Mismatches 210; Indels 151; Gaps 20;
QY 63 AETTWQSDMEWE-----TTVDMTKKQVIFDLSLVKKCLFEVLNTKNIFPGDVNW 114
Db 31 AEKEWELPDSMDNLIEQAPLTVAEKLDQDFLVQWRRVSKA-----PEALF 78
QY 115 FVQHEGKQDQGHCHVLIG-----GKDFSAQAGKWMR-----ROLNVYMSRWLVATCN 162
Db 79 FVQFEKESY-FHLHLIVETTVGKSMVLGRFLSQIRDKLQVITYRGIEPTLPNWFVAVT-- 135
QY 163 VQLTPAERIKLREIADNEWVTLTYKHQTKDYTKCVLFGNMIAYYFLTKKISTSP 222
Db 136 -----KTRNGAGGKNVDECY-----IPNLLPK-----TQP- 163
QY 223 RDGGYFLSDSGWK-TNF-----LKEGERHLVSKLYTDDMRPETVETTTTAQETK 272
Db 164 -----ELQAWNTMEEYISACLNLAERKELVAQHL-----THVSQTQEQN 203
QY 273 RGRIOPTKEVAI---KTLTKE-----LVHKRVTSPEMMQPDYSIEMMAQPGGENLL 323
Db 204 KENLNPNSDAPVIRSKTSARYMELVGLVDRGITSEKQWIOEQDQASYISFNAASNSRSQI 263
QY 324 KNTLEICTLTARTKTAFLILEKATSKLTNFSLPDTRTCRIF-----A 368
Db 264 KAALDNAGKIMALTAKSPDVLVGPAPPA-----DIKTNRIYRILELNGYEPAYAGSV 315
QY 369 FHWGNTVKVCHACCCVNLNQGKGRNTVLPHGPASTGKSIIAQIAQAVGNVGCYNAANVN 428
Db 316 FLGW-----AQKRFGRKNTIWLFGPATGTGNTIAEIAHAHVPYGCNVNTNEN 363
QY 429 FPNDCTNKNLIWVEEAGNFGQVQKALCSGQIRIDQKSGKQIEPTPVIMTNNEN 488
Db 364 FPNDCTNKNLIWVEEAGNFGQVQKALCSGQIRIDQKSGKQIEPTPVIMTNTN 423

QY 429 FPNDCTNKNLIWVEEAGNFGQVQKALCSGQIRIDQKSGKQIEPTPVIMTNNEN 488
Db 364 FPNDCTNKNLIWVEEAGNFGQVQKALCSGQIRIDQKSGKQIEPTPVIMTNTN 423
QY 489 ITVVRIGCEERPEHTQPIRDRMLNIEHLTHLPDGLVLDKNEWPMICAWLVKNGYQSTMA 548
Db 424 MCAVIDGNSITTFEHOQLQDRMPKFEFLTRLEHDFGKVKQEVKPEFRWAQDHVTEVAHE 483
QY 549 SYCAKMG-----KVPDMS--NWAEPKVPTPINLLGSARSPTTP 585
Db 484 FYVRKGGARKRPAPDDADK-SEPKRAC-----SVADPSTS 518
RESULT 5
US-09-807-802A-2
; Sequence 2, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNVN.031USA
; CURRENT FILING DATE: 2002-02-21
; PRIOR FILING DATE: 2002-02-21
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 623
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-2

Query Match 11.9%; Score 433.5; DB 4; Length 623;
Best Local Similarity 27.0%; Pred. No. 5.4e-36;
Matches 157; Conservative 63; Mismatches 210; Indels 151; Gaps 20;
QY 63 AETTWQSDMEWE-----TTVDMTKKQVIFDLSLVKKCLFEVLNTKNIFPGDVNW 114
Db 31 AEKEWELPDSMDNLIEQAPLTVAEKLDQDFLVQWRRVSKA-----PEALF 78
QY 115 FVQHEGKQDQGHCHVLIG-----GKDFSAQAGKWMR-----ROLNVYMSRWLVATCN 162
Db 79 FVQFEKESY-FHLHLIVETTVGKSMVLGRFLSQIRDKLQVITYRGIEPTLPNWFVAVT-- 135
QY 163 VQLTPAERIKLREIADNEWVTLTYKHQTKDYTKCVLFGNMIAYYFLTKKISTSP 222
Db 136 -----KTRNGAGGKNVDECY-----IPNLLPK-----TQP- 163
QY 223 RDGGYFLSDSGWK-TNF-----LKEGERHLVSKLYTDDMRPETVETTTTAQETK 272
Db 164 -----ELQAWNTMEEYISACLNLAERKELVAQHL-----THVSQTQEQN 203
QY 273 RGRIOPTKEVAI---KTLTKE-----LVHKRVTSPEMMQPDYSIEMMAQPGGENLL 323
Db 204 KENLNPNSDAPVIRSKTSARYMELVGLVDRGITSEKQWIOEQDQASYISFNAASNSRSQI 263
QY 324 KNTLEICTLTARTKTAFLILEKATSKLTNFSLPDTRTCRIF-----A 368
Db 264 KAALDNAGKIMALTAKSPDVLVGPAPPA-----DIKTNRIYRILELNGYEPAYAGSV 315
QY 369 FHWGNTVKVCHACCCVNLNQGKGRNTVLPHGPASTGKSIIAQIAQAVGNVGCYNAANVN 428
Db 316 FLGW-----AQKRFGRKNTIWLFGPATGTGNTIAEIAHAHVPYGCNVNTNEN 363
QY 429 FPNDCTNKNLIWVEEAGNFGQVQKALCSGQIRIDQKSGKQIEPTPVIMTNNEN 488
Db 364 FPNDCTNKNLIWVEEAGNFGQVQKALCSGQIRIDQKSGKQIEPTPVIMTNTN 423

QY 489 ITVVRIGCEERBEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVKNGYQSTWA 548
Db 424 MCAVIDGNSTTFHQOPLQDRMFKEFLTRRLEHDFGKVTKEVFRWAQDHVTEVAHE 483

QY 549 SYCAKWG-----KVPDWSNWAEPKVPPTPINLLGSARSPTT 585
Db 484 FYVRKGANKRPAPDDADK-SEPKRACP-----SVADPSTS 518

RESULT 6
US-09-807-802A-5
; Sequence 5, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 623
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-5

Query Match 11.9%; Score 433.5; DB 4; Length 623;
Best Local Similarity 27.0%; Pred. No. 5.4e-36;
Matches 157; Conservative 63; Mismatches 210; Indels 151; Gaps 20;

QY 63 AETTMDQSDMEWE-----TTVDMTKKQVIFDLSLVKGLFEVLNTRKNIFPGDVNW 114
Db 31 AEKEWELPPDSMDLNLIEQAPLTVAEKLQDRPLVQWRVSKA-----PEALF 78

QY 115 FVQHEKGQDQWHCHVLIG-----GKDSQAQGWKR---ROLNVYWRWLVTACN 162
Db 79 FVQFEGESY-FHLHLVETTVGKSMVLGRFLSQIRDKLVQTIYRGIEPTLPNFAVT-- 135

QY 163 VQITPAERIKLREIADNEWVTLTYKHQTKDYTKCVLFGNMIAVYFLTKKISTSP 222
Db 136 -----KTRNGAGGNKVDECY-----IPNYLLPK-----TQP- 163

QY 223 RDGGYFLSSDSGWK-TNF-----LKEGERHLVSKLYTDDMRPETVETVTTTAQETK 272
Db 164 -----ELQAWTNMBEYISACINLAERKELVAQHL-----THVSQTOEQN 203

QY 273 RGRIOQKKEVAI---KTTUKE-----LVHKRVTSPEWMMQPDYSIEMMAQPGGENLL 323
Db 204 KENLNPNSDAPVIRSKTSARYMELVGLVDRGITSEKOWIQEDQASYSIFNAASNSRSQI 263

QY 324 KNTLEICTLTARTKTAFLDILEKETSCLTNFSLPDRTRCRIF-----A 368
Db 264 KAALDNAGKIMALTGAPDVLVGPAPPA-----DIKNRIYRLELNGVEPAGVSV 315

QY 369 FHGWNVYKVCATCCVLNRQGGKRNITVLFHGPASTGKSIIAQIAQAVNGVGCYNAAVNV 428
Db 316 FLGW-----AQKRFGRKNTIWLFGPATTKNTIAEAIAHAVPFYGCNVNNTNEN 363

QY 429 FPNDCNTKNLIWVEAGNFQGVNQFKAICSGQTIIDOKGKSGKQIEPTPVTMTNEN 488
Db 364 FPNDCVCDKMWIWMEEGKMTAKVVEAKAILGSKVRVDQKCKSAQIDPTPVIIVTSNTN 423

QY 489 ITVVRIGCEERBEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVKNGYQSTWA 548
Db 424 MCAVIDGNSTTFHQOPLQDRMFKEFLTRRLEHDFGKVTKEVFRWAQDHVTEVAHE 483

QY 549 SYCAKWG-----KVPDWSNWAEPKVPPTPINLLGSARSPTT 585
Db 484 FYVRKGANKRPAPDDADK-SEPKRACP-----SVADPSTS 518

RESULT 7
US-09-532-594B-8
; Sequence 8, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; OTHER INFORMATION: synthetic construct
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 Rep protein 40
US-09-532-594B-8

Query Match 11.8%; Score 430.5; DB 4; Length 312;
Best Local Similarity 35.6%; Pred. No. 3.2e-36;
Matches 110; Conservative 45; Mismatches 139; Indels 15; Gaps 5;

QY 291 LVHKRVTSPEWMMQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAE 350
Db 7 LVDRGITSEKOWIQEDQASYSIFNAASNSRSQIKAALDNASKIMSLTKTAPDVLVGNPP 66

QY 351 SKLTNPSLPDRTRCRIFAFHGWNVYKVCATCCVLNRQGGKRNITVLFHGPASTGKSIIAQ 410
Db 67 EBISS-----NRIYRLELNGYDPOVAAASVFLGWAQKFGKNTIWLFGPATTKNTIAE 121

QY 411 AIAQAVNGVGCYNAAVNVFPNDCTNKNLIWVEAGNFQGVNQFKAICSGQTIIDOKG 470
Db 122 AIAHAVPFYGCNVNNTNENFPNDCTNKNLIWVEAGNFQGVNQFKAICSGQTIIDOKG 181

QY 471 KGSQKQIEPTPVTMTNENITVVRIGCEERBEHTQPIRDRMLNIHLTHLPGDFGLVDKNE 530
Db 182 KSSAQIDPTPVIIVTSNTNMCVIDGNSTTFHQOPLQDRMFKEFLTRRLEHDFGKVTKE 241

QY 531 WPMICAWLVKNGYQSTMASYCAKWG--KVPDWSNWAEPKVPPTPINLLGSARSPTTTPK 587
Db 242 VKDFFRWASDHVTEVTHEFYVRKGARKPAPDNADADISEPKRACP-----SVAQPS 296

QY 588 STPLSQNYA 596
Db 297 EAPV--DYA 303

RESULT 8
US-09-532-594B-9
; Sequence 9, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22


```
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6469524e =
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 Rep protein 52
US-09-532-594B-9

Query Match      11.8%; Score 430.5; DB 4; Length 399;
Best Local Similarity 35.0%; Pred. No. 4.9e-36;
Matches 110; Conservative 45; Mismatches 139; Indels 15; Gaps 5;

QY 291 LVHKRVTSPEDNMMPDSDYIEMMAQPGGENLLKNTLEICTLTLARTKTAFDLILEKAET 350
Db 7 LVDRGITSEKQWIOEQDQASYISFNAASRSQIKAAALDNAGKIMALTAKSAPDYLVGPP 66

QY 351 SKLTNFSLPDTRTCRIF-----AFHGMWYVVKVCHAICCVLNROGKENTV 395
Db 67 A-----DIKTNRIYRIELNGYEPAYAGSVFLGW-----AQKRFGRKNTI 106

QY 396 LFHGPASTGKSIIAQIAQAVGNVGCYNAAVNFNFNDCTNKNLIWVEEAGNFGQQVNOF 455
Db 107 WLFPGPATTGKTNIAEIAHAHVPYGCNVNTNENFPFNDQVCKVWVWEEGKMTAKVESA 166

QY 456 KAICSGQTIRIDQKSGKQIEPTPVIMTNTNENITVVRIGCEERPEHTQPIRDMNLHL 515
Db 167 KAILGSKVRVDQKCKSSAQIDPTPVIVTSNTNCAVIDGNSNTHFHHQQLQDRMFKEFEL 226

QY 516 THHLPDGLVDKNWPMICAMLVKNGYQSTWASCAKMG-----KVPDWSENNAEPKVPT 571
Db 227 TRLEHDFGKVTQKEVKEFFRWAQDHVTEVAHEFFYVRKGGANKRPAPDDADK--SEPKRAC 285

QY 572 PINLLGSARSPFTT 585
Db 286 P-----SVADPSTS 294

RESULT 10
US-09-807-802A-9
; Sequence 9, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-9

Query Match      11.7%; Score 425.5; DB 4; Length 399;
Best Local Similarity 35.0%; Pred. No. 1.7e-35;
Matches 110; Conservative 37; Mismatches 122; Indels 45; Gaps 6;

QY 291 LVHKRVTSPEDNMMPDSDYIEMMAQPGGENLLKNTLEICTLTLARTKTAFDLILEKAET 350
Db 7 LVDRGITSEKQWIOEQDQASYISFNAASRSQIKAAALDNAGKIMALTAKSAPDYLVGPP 66

QY 351 SKLTNFSLPDTRTCRIF-----AFHGMWYVVKVCHAICCVLNROGKENTV 395
Db 67 A-----DIKTNRIYRIELNGYEPAYAGSVFLGW-----AQKRFGRKNTI 106

QY 396 LFHGPASTGKSIIAQIAQAVGNVGCYNAAVNFNFNDCTNKNLIWVEEAGNFGQQVNOF 455
Db 107 WLFPGPATTGKTNIAEIAHAHVPYGCNVNTNENFPFNDQVCKVWVWEEGKMTAKVESA 166

QY 456 KAICSGQTIRIDQKSGKQIEPTPVIMTNTNENITVVRIGCEERPEHTQPIRDMNLHL 515
Db 167 KAILGSKVRVDQKCKSSAQIDPTPVIVTSNTNCAVIDGNSNTHFHHQQLQDRMFKEFEL 226

QY 516 THHLPDGLVDKNWPMICAMLVKNGYQSTWASCAKMG-----KVPDWSENNAEPKVPT 571
Db 227 TRLEHDFGKVTQKEVKEFFRWAQDHVTEVAHEFFYVRKGGANKRPAPDDADK--SEPKRAC 285

; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6469524e =
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 Rep protein 52
US-09-532-594B-9

Query Match      11.8%; Score 430.5; DB 4; Length 399;
Best Local Similarity 35.0%; Pred. No. 4.9e-36;
Matches 110; Conservative 45; Mismatches 139; Indels 15; Gaps 5;

QY 291 LVHKRVTSPEDNMMPDSDYIEMMAQPGGENLLKNTLEICTLTLARTKTAFDLILEKAET 350
Db 7 LVDRGITSEKQWIOEQDQASYISFNAASRSQIKAAALDNAGKIMALTAKSAPDYLVGPP 66

QY 351 SKLTNFSLPDTRTCRIFAFHGMWYVVKVCHAICCVLNROGKENTVLFHGPASTGKSIIAQ 410
Db 67 EDISS-----NRIYRIELNGYDPQYAAASVFLGWAQKTKGRNTIWLFGPATTGKTNI 121

QY 411 ATAAVGNVGCYNAAVNFNFNDCTNKNLIWVEEAGNFGQQVNOFKAICSGQTIRIDQKG 470
Db 122 ATAAHVPYGCNVNTNENFPFNDQVCKVWVWEEGKMTAKVESA KAILGSKVRVDQK 181

QY 471 KGSQKLEPTPVIMTNTNENITVVRIGCEERPEHTQPIRDMNLHLTHHLPDGLVDKNE 530
Db 182 KSSAQIDPTPVIVTSNTNCAVIDGNSNTHFHHQQLQDRMFKEFELTRLEHDFGKVTQKE 241

QY 531 WPMICAWLVKNGYQSTWASCAKMG--KVPDWS--NWAEPKVPTPINLLGSARSPFTTPK 587
Db 242 VKDFFRWASDHVTEVTHEFVVRKGGARKRPAPDDADISEPKRACP-----SVAQPSTSDA 296

QY 588 STPLSQNYA 596
Db 297 EAPV--DYA 303

RESULT 9
US-09-807-802A-11
; Sequence 11, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 322
; TYPE: PRT
; ORGANISM: AAV-1
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (943)..(944)
; OTHER INFORMATION: minor splice site
US-09-807-802A-11

Query Match      11.7%; Score 425.5; DB 4; Length 322;
Best Local Similarity 35.0%; Pred. No. 1.1e-35;
Matches 110; Conservative 37; Mismatches 122; Indels 45; Gaps 6;
```

QY 572 PINLLGARSPTT 585
Db 286 P-----SVADPSTS 294

RESULT 11
US-09-171-461-3
; Sequence 3, Application US/09171461
; Patent No. 6335016
; GENERAL INFORMATION:
; APPLICANT: Baker, Adam
; APPLICANT: Cotten, Matthew
; APPLICANT: Chioccia, Susanna
; APPLICANT: Kurzbauer, Robert
; APPLICANT: Schaffner, Gotthold
; TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus
; FILE REFERENCE: 0652.180000
; CURRENT APPLICATION NUMBER: US/09/171,461
; CURRENT FILING DATE: 1999-01-12
; EARLIER APPLICATION NUMBER: PCT/EP97/01944
; EARLIER FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 276
; TYPE: PRT
; ORGANISM: CELO Virus
; FEATURE:
; OTHER INFORMATION: Position: 1999..2829 /note=ORF2
US-09-171-461-3

Query Match 4.0%; Score 144; DB 3; Length 276;
Best Local Similarity 24.1%; Pred. No. 3.6e-06;
Matches 64; Conservative 41; Mismatches 133; Indels 28; Gaps 9;
QY 286 TTLKELVHKRVTSPEDEMMQPDYS--IEMMAQPGGENLLKNTLEICTTLTARTKTAFDL 343
Db 10 TLVHALIDRGVVSREQQWVDPAQYQFYHRSKQGFK--VRHLRDVIRHMCWSRTLLDY 67
QY 344 ILEKAETSKLTNFSLPD----TRTCRIFAFHGWNTYKVKCHAI CCVLNRQGGKRNITVLFHG 399
Db 68 MSSASTPS-----PDDVLRNPLYQLLCCNGYNPAVVGTTALIRWAGHQ--SNRNTVWIRG 119
QY 400 PASTGKSIITAAQIAAAGVNGVCYNAANVPNDCTNKNLIWVEEAGNFGQOVNQFKAIC 459
Db 120 TPMGAPYLAQAIAYCSPLVGSVDWRNKSNNPFEGCPDLSLVFWDGGYVYDCCVGLVKQVF 179
QY 460 SGQTIRIDQK----KGSQKIEPTVIMTNNITVVRIGCEE-RPEHTQPIRDRMLNTH 514
Db 180 RGEHVILPPEGLRGNPCSELFRTPVLMYSQADICWTRLRSGLSAEHAVALRDCMYLIR 239
QY 515 LTHLPGDF---GLVDKNEWPMICAW 537
Db 240 LTE----DFDCAGGISCADVKQFVAW 261

RESULT 12
US-09-970-711-3
; Sequence 3, Application US/09970711
; Patent No. 6773709
; GENERAL INFORMATION:
; APPLICANT: Baker, Adam
; APPLICANT: Cotten, Matthew
; APPLICANT: Chioccia, Susanna
; APPLICANT: Kurzbauer, Robert
; APPLICANT: Schaffner, Gotthold
; TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus
; FILE REFERENCE: 0652.180001
; CURRENT APPLICATION NUMBER: US/09/970,711
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/171,461
; PRIOR FILING DATE: 1999-01-12

; PRIOR APPLICATION NUMBER: PCT/EP97/01944
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 276
; TYPE: PRT
; ORGANISM: CELO Virus
; FEATURE:
; OTHER INFORMATION: Position: 1999..2829 /note=ORF2
US-09-970-711-3

Query Match 4.0%; Score 144; DB 4; Length 276;
Best Local Similarity 24.1%; Pred. No. 3.6e-06;
Matches 64; Conservative 41; Mismatches 133; Indels 28; Gaps 9;
QY 286 TTLKELVHKRVTSPEDEMMQPDYS--IEMMAQPGGENLLKNTLEICTTLTARTKTAFDL 343
Db 10 TLVHALIDRGVVSREQQWVDPAQYQFYHRSKQGFK--VRHLRDVIRHMCWSRTLLDY 67
QY 344 ILEKAETSKLTNFSLPD----TRTCRIFAFHGWNTYKVKCHAI CCVLNRQGGKRNITVLFHG 399
Db 68 MSSASTPS-----PDDVLRNPLYQLLCCNGYNPAVVGTTALIRWAGHQ--SNRNTVWIRG 119
QY 400 PASTGKSIITAAQIAAAGVNGVCYNAANVPNDCTNKNLIWVEEAGNFGQOVNQFKAIC 459
Db 120 TPMGAPYLAQAIAYCSPLVGSVDWRNKSNNPFEGCPDLSLVFWDGGYVYDCCVGLVKQVF 179
QY 460 SGQTIRIDQK----KGSQKIEPTVIMTNNITVVRIGCEE-RPEHTQPIRDRMLNTH 514
Db 180 RGEHVILPPEGLRGNPCSELFRTPVLMYSQADICWTRLRSGLSAEHAVALRDCMYLIR 239
QY 515 LTHLPGDF---GLVDKNEWPMICAW 537
Db 240 LTE----DFDCAGGISCADVKQFVAW 261

RESULT 13
US-09-300-909-17
; Sequence 17, Application US/09300909
; Patent No. 6306580
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PREPARATION OF HUMAN PAPILLOMAVIRUS E1 HAVING
; TITLE OF INVENTION: HELICASE ACTIVITY AND METHOD THEREFOR
; NUMBER OF SEQUENCES: 27
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/300,909
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/083,942
; FILING DATE: 01-MAY-1998
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 647 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-300-909-17

Query Match 3.4%; Score 122.5; DB 3; Length 647;
Best Local Similarity 19.7%; Pred. No. 0.0031;
Matches 110; Conservative 77; Mismatches 173; Indels 197; Gaps 26;
QY 39 GKDIGWNSYKKEIQEDELK-SLQGAETTWQDSE----DMEWETT-VDEMTKQVIFDS 92
Db 100 GKVGGRNT-RGTLQELSLNVSSTQATQTVYSPDSGYGNNEVETAEVEEVT----- 149

QY 93 LVKKCLFEVLNTKNIFFGDVNVFVQHEWKGQDQGHCHVLIGGKDFSOAQGWRRQLNVY 152
Db 150 -----VATNTNGDAEGHGSVREE-----CSSVDSATDSNQDPK----- 185
QY 153 WSRWLVTACNVQLTPAERIKLREIAEDNEWVTLTYKHQT-----KKDYTKC 200
Db 186 -----SPTAQIKLL-LQSNKKAAMLT-QFKETYGLSFTDLVRFKSKTTC 230
QY 201 VLFGNMIAYYFLTKKISTSPRGGYFLSSDSGWKTNFLKEGERHLVSK--LYTDDMRP 258
Db 231 T---DWAAIFGVHPTIA-----EGFKTLINKYALYTH--- 260
QY 259 ETVEITVTTAETKRGRIOT-----KKEVAIKTTLKELVHKRVTSPEDEMMQPDYSY 310
Db 261 -----IQSLDTKQGVLLMLIRYTCGRNVRTVGKLSLTH-----VPESCMLLLEPPKL 309
QY 311 IEMMA-----OPGENLLKNTLE-----ICTLTAR---TKTAFDL--ILEKATSCLT-- 354
Db 310 RSPVAALYWRIGISNISVVTGTPWIORLTVIQHIGIDDSVFDLSDMVQWAFDNEYTDE 369
QY 355 -----NFSLPDTRTCRIFAFHGMN-----YVKVCHAIC-----CVL 385
Db 370 SDIAFNAYMLADCSNAAFLKSNCOAKYKDCATWCKHKYKRAQKQMSQWIKERCCK 429
QY 386 NROGG-----KRNVTLPFGPASTGCKSIIAQIAQ 414
Db 430 CDEGGDWRPIQVFLRYOGIEFISFLCALKEFLKGTGPKQNCIVYGANTGKSHFCMSLMH 489
QY 415 AV-GNVGCYNAANVNPFCNTKNIWVEAGN--FGQVNOFKALCSGOTIRIDOKG 471
Db 490 FLOGTVISVNSTSHFWLSEPLADAKLAMLDDATGTCWCFDNTMRNALDGYAISLDRKYK 549
QY 472 GSKQIEPTVIMTNTEN 488
Db 550 SLLQMKCPPLITSNTN 566

RESULT 14
US-09-107-532A-6086
; Sequence 6086, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneka
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6086:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 589 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...589
; SEQUENCE DESCRIPTION: SEQ ID NO: 6086:
US-09-107-532A-6086

Query Match 3.1%; Score 113; DB 4; Length 589;
Best Local Similarity 19.9%; Pred. No. 0.026;
Matches 94; Conservative 60; Mismatches 150; Indels 168; Gaps 23;

QY 162 NVQLTPAERIKUREIAEDNEWVTLT-YKHQTKDYTKCVLFGNMIAYYFLTKKISTS 220
Db 142 NIKLRNATTKAIRHYLDDNDFLDIETPYLGKSTPEG-----ARDYLVPSRV--- 187
QY 221 PRDGGYFLS---SDSGWKNFLKEG--ERHLVSKLYTD-----DMRPE---TVETTVT 266
Db 188 ---HAGHFYALPQSPQLFKQLLMNAGFDRIYQIVRCFRDEDLGRQPEFTQVDIETFL 244
QY 267 TAQETKRGRIOTKKEVAIKTTLKELVHKRVTSPEDEMMQPDYSIEMMAQPGGENLLKNT 326
Db 245 TAEE-----IQYTEGLIAKWKVEGVIEVTLFPFM-----TYDEAMARYGSD----- 288
QY 327 LEICTLTLARTKTAFDL-ILEKAETSCLTNFSLPDRTRTCRIF--AFHGMNVYKVCCHAIC 383
Db 289 -----KPDTRFDMELIDLSDTVKEVEF-----KVFQMALENGGVVKALNA--- 328
QY 384 VLNRQGGKRNVTLPFGPASTGKSIIAQIAQAVGNVGCYNAANVNPFCNTKNIW--- 441
Db 329 -----KGADRYSRKMDQLGQVYVGQFGA-----KGLAWLK 359
QY 442 VEEAGNFQGVNOFKALCSGQTRIDQKKGSKQIEPTVIM--TTNENITVVRIGCEER 499
Db 360 VEEDGLKG-PIAKFMGEATEAIK-----ATDAKPGDLLMFGADKSEIVAALGA--- 408
QY 500 PEHTQPIRDRMLNLIHTHLPGDFGLVDKNWPMICAWLVKNGYQSTMASYCAKWKVPD 559
Db 409 -----IRTR-----LGKELGLIDSKFNFL----- 428
QY 560 WSENWAEPKVPTPINLLGSARSPFTTPKSTPLSONYALTPLASDLDELALPEP 611
Db 429 WVTDMWPQFEFSEEGRYVSAHHPTMPKE-----BDIPLAEDP 467

RESULT 15
US-09-134-001C-2897
; Sequence 2897, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; ENTEROCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ IDS NOS: 5674
; SEQ ID NO 2897
; LENGTH: 564
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:41:01 ; Search time 90.6667 Seconds
(without alignments)
2677.792 Million cell updates/sec

Title: US-10-069-056-6
Perfect score: 3638
Sequence: 1 MAGNAYSEVLGATNWLKEK.....RACFGAEPLKDFSEPLNLD 672

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues
Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.ppt.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.ppt.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.ppt.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.ppt.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.ppt.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.ppt.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.ppt.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.ppt.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.ppt.*
11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.ppt.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.ppt.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.ppt.*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.ppt.*
15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.ppt.*
16: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.ppt.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.ppt.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.ppt.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.ppt.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.ppt.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	455.5	12.5	610	9	US-09-792-630-21
2	455.5	12.5	610	10	US-09-953-351-21
3	455.5	12.5	610	13	US-10-080-376-21
4	455.5	12.5	610	14	US-10-082-671-27
5	455.5	12.5	610	14	US-10-097-100-21
6	455.5	12.5	610	14	US-10-023-208-21
7	455.5	12.5	610	14	US-10-375-192-7
8	452.5	12.4	626	9	US-09-792-630-29
9	452.5	12.4	626	13	US-09-953-351-29
10	452.5	12.4	626	13	US-10-080-376-29
11	452.5	12.4	626	14	US-10-082-671-35
12	452.5	12.4	626	14	US-10-097-100-29
13	452.5	12.4	626	14	US-10-023-208-29

14	452	12.4	627	9	US-09-792-630-25	Sequence 25, Appl
15	452	12.4	627	10	US-09-953-351-25	Sequence 25, Appl
16	452	12.4	627	13	US-10-080-376-25	Sequence 25, Appl
17	452	12.4	627	14	US-10-082-671-31	Sequence 31, Appl
18	452	12.4	627	14	US-10-097-100-25	Sequence 25, Appl
19	452	12.4	627	14	US-10-023-208-25	Sequence 25, Appl
20	445.5	12.2	625	14	US-10-423-704A-3	Sequence 3, Appl
21	437.5	12.0	537	10	US-09-254-747-10	Sequence 10, Appl
22	437.5	12.0	623	9	US-09-792-630-5	Sequence 5, Appl
23	437.5	12.0	623	10	US-09-953-351-5	Sequence 2, Appl
24	437.5	12.0	623	10	US-09-254-747-11	Sequence 11, Appl
25	437.5	12.0	623	13	US-10-080-376-5	Sequence 5, Appl
26	437.5	12.0	623	13	US-10-082-671-11	Sequence 11, Appl
27	437.5	12.0	623	14	US-10-097-100-5	Sequence 5, Appl
28	437.5	12.0	623	14	US-10-023-208-5	Sequence 5, Appl
29	437.5	12.0	623	14	US-10-375-192-5	Sequence 5, Appl
30	437.5	12.0	623	9	US-09-792-630-13	Sequence 13, Appl
31	436.5	12.0	623	10	US-09-953-351-13	Sequence 13, Appl
32	436.5	12.0	623	13	US-10-080-376-13	Sequence 13, Appl
33	436.5	12.0	623	14	US-10-082-671-19	Sequence 19, Appl
34	436.5	12.0	623	14	US-10-097-100-13	Sequence 13, Appl
35	436.5	12.0	623	14	US-10-023-208-13	Sequence 13, Appl
36	436.5	12.0	623	14	US-10-375-192-2	Sequence 2, Appl
37	436.5	12.0	623	14	US-10-696-261-7	Sequence 7, Appl
38	433.5	11.9	546	15	US-10-696-282-7	Sequence 7, Appl
39	433.5	11.9	546	15	US-10-696-980-7	Sequence 7, Appl
40	433.5	11.9	546	15	US-10-696-980-7	Sequence 7, Appl
41	433.5	11.9	623	9	US-09-792-630-11	Sequence 11, Appl
42	433.5	11.9	623	10	US-09-953-351-11	Sequence 11, Appl
43	433.5	11.9	623	13	US-10-080-376-11	Sequence 11, Appl
44	433.5	11.9	623	14	US-10-082-671-17	Sequence 17, Appl
45	433.5	11.9	623	14	US-10-097-100-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-09-792-630-21
; Sequence 21, Application US/09792630
; Patent No. US20020168640A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES
; FILE REFERENCE: A-70295/RT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/792,630
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 21
; LENGTH: 610
; TYPE: PRT
; ORGANISM: adeno-associated virus 5
US-09-792-630-21

Query Match	12.5%	Score 455.5;	DB 9;	Length 610;
Best Local Similarity	27.8%	Pred. No. 1.2e-30;		
Matches 162;	Conservative 78;	Mismatches 217;	Indels 125;	Gaps 23;
QY	67	WQSEDMENETTVDENETKQVIFDLSLVKKCLPEVLNTKNI	PFQDYNWVFOHEWGRDQGW	126
Db	35	WELPPESDLNLTAVE--QPQLTVADRIIRVFLYE---WNKFSKQESKFFVFKGSEY-F	88	
QY	127	HCHVLIGGKDF-SQAQGWRRRLNLYVWSRWLTACNVOLTPAERIKLREI-AEDNEWVT	184	
Db	89	HLHLTVETSGISWLVGRY-----VSQIRALV---KVVFGIEQINDWA	132	
QY	185	LITYGHKQTKDYTKCVLFENMIAYFLTKKISTSPRDRGGYFLSDSGWK-TNF----	239	
Db	133	I-----TKVKGKGANVDSGYIPAYLLPK-----VQPELQWMTNLDEYK	173	
QY	240	-----LKEGERHLVSKLYTDDMRPETVTTVTAQETKGR-----IQTKKEVAIKTT	287	

Db 174 LAALNLEERKRLVAQFLAESSQR-----SQEAASQREFSADPVIKSKTSQKYMAL 223
Qy 288 LKELVHKVTSPEDEMMQWQDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILLEK 347
Db 224 VNWLVHEGITSSEKQWIOENQESYLSFNSTGNSRSQIKAAALDNATKIMSLTSAVDY---- 279
Qy 348 AETSKLTNFSLPD-----TRTCRIFAFHGWN--YV-KVCHAI CCVLRQGGKRNVTLPFHGP 400
Db 280 -----LVGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGCW---QSFNKRNTVWLYGP 331
Qy 401 ASTGKSIIAQAIAQAVGNCVNAANVPFPNDCTNKNLIWVEEAGNFQOQVNOFKAICS 460
Db 332 ATTGKTINIAEAIAHTVPFYGCNVNENFPFNDCTNKNLIWVEEAGNFQOQVNOFKAICS 391
Qy 461 GOTIRIDOKGKSGKOIEPTPVIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPL 520
Db 392 GSKVRVDOKCKSSQVIDSTFPIVTSNTNMCVVVDGNSSTTFEHQOQPLEDRMFKFELTKRLP 451
Qy 521 GDFGLVDKNEWPMICAWLVKNGYQSTWASACAKWGVDPDWE-----NWAEPKVP 570
Db 452 PDFGKITKQEVKDFFAW-----AKVQVPVTHFEKVPRELAGTKGAESLK 497
Qy 571 TPI-----NLLGSARSPT--TPKSTPLSQNYA-LTPL 600
Db 498 RPLGVDVTNTSYKSLERARLSFVPTPRSSDVTVDPAPLRPL 539

RESULT 2

US-09-953-351-21
; Sequence 21, Application US/09953351
; Publication No. US2003036643A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Melander, Christian
; APPLICANT: Liu, Hong-Xiang
; APPLICANT: Jin, Cheng He
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FUSION I
; FILE REFERENCE: A-70814/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/953,351
; CURRENT FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/232,960
; PRIOR FILING DATE: 2000-09-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 610
; TYPE: PRT
; ORGANISM: adeno-associated virus 5
US-09-953-351-21

Query Match 12.5%; Score 455.5; DB 10; Length 610;
Best Local Similarity 27.8%; Pred. No. 1.2e-30;
Matches 162; Conservative 78; Mismatches 217; Indels 125; Gaps 23;
Qy 67 WQSEDMEWETTVDKTKQVIFDLSLVKCKLFEVLNTKNIFFPGDVNMFVQHEWGKDOGW 126
Db 35 WELPPESDLNLTVE--QPQLTVADRIRRVFLYE---WNKFSKQESKFFVQFEKGESEY-F 88
Qy 127 HCHVLIGKDF-SQAGKWRRLQVYWSRWLVTCNVQLTPAERIKUREI-AEDNEWT 184
Db 89 HLHTLVETSGISSMVLGRY-----VSQIRALV---KVVFQGIPEQINDWVA 132
Qy 185 LLTYKHOTKDYTKCVLFGNMIAYYFLTKKISTSPRDGGYFLSSDSGWK-TNF---- 239
Db 133 I-----TKVKGKANGKVDVSGIIPAYLLPK-----VOPELQAWNTLDEYK 173
Qy 240 -----LKEGERHLVSKLYTDDMRPETVETVTTAQTAKRGR-----IOTKKEVAIKTT 287
Db 174 LAALNLEERKRLVAQFLAESSQR-----SQEAASQREFSADPVIKSKTSQKYMAL 223
Qy 288 LKELVHKVTSPEDEMMQWQDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILLEK 347
Db 174 LAALNLEERKRLVAQFLAESSQR-----SQEAASQREFSADPVIKSKTSQKYMAL 223
Qy 288 LKELVHKVTSPEDEMMQWQDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILLEK 347

Db 224 VNWLVHEGITSSEKQWIOENQESYLSFNSTGNSRSQIKAAALDNATKIMSLTSAVDY---- 279
Qy 348 AETSKLTNFSLPD-----TRTCRIFAFHGWN--YV-KVCHAI CCVLRQGGKRNVTLPFHGP 400
Db 280 -----LVGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGCW---QSFNKRNTVWLYGP 331
Qy 401 ASTGKSIIAQAIAQAVGNCVNAANVPFPNDCTNKNLIWVEEAGNFQOQVNOFKAICS 460
Db 332 ATTGKTINIAEAIAHTVPFYGCNVNENFPFNDCTNKNLIWVEEAGNFQOQVNOFKAICS 391
Qy 461 GOTIRIDOKGKSGKOIEPTPVIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPL 520
Db 392 GSKVRVDOKCKSSQVIDSTFPIVTSNTNMCVVVDGNSSTTFEHQOQPLEDRMFKFELTKRLP 451
Qy 521 GDFGLVDKNEWPMICAWLVKNGYQSTWASACAKWGVDPDWE-----NWAEPKVP 570
Db 452 PDFGKITKQEVKDFFAW-----AKVQVPVTHFEKVPRELAGTKGAESLK 497
Qy 571 TPI-----NLLGSARSPT--TPKSTPLSQNYA-LTPL 600
Db 498 RPLGVDVTNTSYKSLERARLSFVPTPRSSDVTVDPAPLRPL 539

RESULT 3

US-10-080-376-21
; Sequence 21, Application US/10080376
; Publication No. US20020172968A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Dahiyat, Bassil I.
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES
; FILE REFERENCE: A-70295-2/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/080,376
; CURRENT FILING DATE: 2000-02-19
; PRIOR APPLICATION NUMBER: US 09/792,630
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 610
; TYPE: PRT
; ORGANISM: adeno-associated virus 5
US-10-080-376-21

Query Match 12.5%; Score 455.5; DB 13; Length 610;
Best Local Similarity 27.8%; Pred. No. 1.2e-30;
Matches 162; Conservative 78; Mismatches 217; Indels 125; Gaps 23;
Qy 67 WQSEDMEWETTVDKTKQVIFDLSLVKCKLFEVLNTKNIFFPGDVNMFVQHEWGKDOGW 126
Db 35 WELPPESDLNLTVE--QPQLTVADRIRRVFLYE---WNKFSKQESKFFVQFEKGESEY-F 88
Qy 127 HCHVLIGKDF-SQAGKWRRLQVYWSRWLVTCNVQLTPAERIKUREI-AEDNEWT 184
Db 89 HLHTLVETSGISSMVLGRY-----VSQIRALV---KVVFQGIPEQINDWVA 132
Qy 185 LLTYKHOTKDYTKCVLFGNMIAYYFLTKKISTSPRDGGYFLSSDSGWK-TNF---- 239
Db 133 I-----TKVKGKANGKVDVSGIIPAYLLPK-----VOPELQAWNTLDEYK 173
Qy 240 -----LKEGERHLVSKLYTDDMRPETVETVTTAQTAKRGR-----IOTKKEVAIKTT 287
Db 174 LAALNLEERKRLVAQFLAESSQR-----SQEAASQREFSADPVIKSKTSQKYMAL 223
Qy 288 LKELVHKVTSPEDEMMQWQDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILLEK 347
Db 224 VNWLVHEGITSSEKQWIOENQESYLSFNSTGNSRSQIKAAALDNATKIMSLTSAVDY---- 279
Qy 348 AETSKLTNFSLPD-----TRTCRIFAFHGWN--YV-KVCHAI CCVLRQGGKRNVTLPFHGP 400
Db 280 -----LVGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGCW---QSFNKRNTVWLYGP 331
Qy 401 ASTGKSIIAQAIAQAVGNCVNAANVPFPNDCTNKNLIWVEEAGNFQOQVNOFKAICS 460


```
Db 332 ATTGKTNIAEAIHTVFFYGCNVWNTNFFNDCVDMKLIWEEGKMTNKVVSAAKAILG 391
Qy 461 GQIRIDQKQKSGKQIEPTPVIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPL 520
Db 392 GSKVRVDQKCKSSVQIDSTPVIIVTSNTNMCVVVDGNSSTTFEHQHPLEDRMFKELTKRLP 451
Qy 521 GDFGLVDKNWPMICAWLVKNGYQSTWASYCAKWKGVDPWSE-----NWAEPKVP 570
Db 452 PDFGKITQKQEVKDFFAW-----AKVNOVPVTHEFKVPRELATGKAESKSLK 497
Qy 571 TPI-----NLLGSARSPT--TPKSTPLSONYA-LTPL 600
Db 498 RPLGVDNTSYKSLEKRLARLSFVPETPRSSDVTVDPAIRPL 539

RESULT 4
US-10-082-671-27
; Sequence 27, Application US/10082671
; Publication No. US20030049647A1
; GENERAL INFORMATION:
; APPLICANT: DAHIYAT, BASSIL
; APPLICANT: LI, MIN
; TITLE OF INVENTION: USE OF NUCLEIC ACID LIBRARIES TO CREATE TOXICOLOGICAL
; FILE REFERENCE: XEN/001
; CURRENT APPLICATION NUMBER: US/10/082, 671
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/270, 781
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 27
; LENGTH: 610
; TYPE: PRT
; ORGANISM: adeno-associated virus 5
US-10-082-671-27

Query Match 12.5%; Score 455.5; DB 14; Length 610;
Best Local Similarity 27.8%; Pred. No. 1.2e-30;
Matches 162; Conservative 78; Mismatches 217; Indels 125; Gaps 23;

Qy 67 WDQSDMEWETTVDKTKQVFIQFDSLVKKCLPEVLTKNIFPGDVNMFVQHEWKGDOGW 126
Db 35 WELPPESDLNLTLVE--QPOLTVDADRIRRVFLYE--WNKFSQKESKFFVQPEKSGSEY-F 88
Qy 127 HCHVLIGGKDF-SQAQKWMRRQLNVYWSRWLTACNVOLTPAERIKLREI-AEDNEWVT 184
Db 89 HLHVLVETSGISSWVLGRY-----VSQIRAOQLV---KVVFQIEQINDWA 132
Qy 185 LLYKHQTKQDYTKCVLFGNMIAYFLTKKLISTSPRRDGGYFLSSDSGWK-TNF---- 239
Db 133 I-----TKVKGKANKVDVSGYIPAYLLPK-----VQPELOWAWTNLDEYK 173
Qy 240 -----LKEGERHLVSKLYTDDMRPETVETVTTTAQETKGR-----IOTKEVAIKTT 287
Db 174 LAALNLEERKRLVAQFLAESSOR-----SQEAASQREPSADPVIKSKTSQKYMAL 223
Qy 288 LKELVHKRVTSPEWMMQPDYSIENMAOPGGENLKNLTLEICTLTLARTKTAFLDILEK 347
Db 224 VNWLVHGHTSEKQWIQENQESYLSFNSTGNSRSQIKAAALDNATKIMSLTKSAVDY---- 279
Qy 348 AETSKLNTSLPD-----TRTCRIFAFHGMN--YV-KVCHAI CCVLRQGGKRNVTLPFGP 400
Db 280 -----LVGSSVPEIDISKRIWQIFEMNGYDPAVAGSILYGCW---QRSFNKRNVTWLYGP 331
Qy 401 ASTGKSIIAQAATAQAGVNGVCYNAANVFPNDCTNKNLIWEEAGNFGQGVNQFKAICS 460
Db 332 ATTGKTNIAEAIHTVFFYGCNVWNTNFFNDCVDMKLIWEEGKMTNKVVSAAKAILG 391
Qy 461 GQIRIDQKQKSGKQIEPTPVIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPL 520
Db 392 GSKVRVDQKCKSSVQIDSTPVIIVTSNTNMCVVVDGNSSTTFEHQHPLEDRMFKELTKRLP 451
Qy 521 GDFGLVDKNWPMICAWLVKNGYQSTWASYCAKWKGVDPWSE-----NWAEPKVP 570
Db 452 PDFGKITQKQEVKDFFAW-----AKVNOVPVTHEFKVPRELATGKAESKSLK 497
```

```
Qy 521 GDFGLVDKNWPMICAWLVKNGYQSTWASYCAKWKGVDPWSE-----NWAEPKVP 570
Db 452 PDFGKITQKQEVKDFFAW-----AKVNOVPVTHEFKVPRELATGKAESKSLK 497
Qy 571 TPI-----NLLGSARSPT--TPKSTPLSONYA-LTPL 600
Db 498 RPLGVDNTSYKSLEKRLARLSFVPETPRSSDVTVDPAIRPL 539

RESULT 5
US-10-097-100-21
; Sequence 21, Application US/10097100
; Publication No. US20030068649A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Melander, Christian
; APPLICANT: Liu, Hong-Xiang
; APPLICANT: Jin, Cheng He
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FUSION LI
; FILE REFERENCE: A-70814/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/097,100
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/953,351
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/232,960
; PRIOR FILING DATE: 2000-09-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 21
; LENGTH: 610
; TYPE: PRT
; ORGANISM: adeno-associated virus 5
US-10-097-100-21

Query Match 12.5%; Score 455.5; DB 14; Length 610;
Best Local Similarity 27.8%; Pred. No. 1.2e-30;
Matches 162; Conservative 78; Mismatches 217; Indels 125; Gaps 23;

Qy 67 WDQSDMEWETTVDKTKQVFIQFDSLVKKCLPEVLTKNIFPGDVNMFVQHEWKGDOGW 126
Db 35 WELPPESDLNLTLVE--QPOLTVDADRIRRVFLYE--WNKFSQKESKFFVQPEKSGSEY-F 88
Qy 127 HCHVLIGGKDF-SQAQKWMRRQLNVYWSRWLTACNVOLTPAERIKLREI-AEDNEWVT 184
Db 89 HLHVLVETSGISSWVLGRY-----VSQIRAOQLV---KVVFQIEQINDWA 132
Qy 185 LLYKHQTKQDYTKCVLFGNMIAYFLTKKLISTSPRRDGGYFLSSDSGWK-TNF---- 239
Db 133 I-----TKVKGKANKVDVSGYIPAYLLPK-----VQPELOWAWTNLDEYK 173
Qy 240 -----LKEGERHLVSKLYTDDMRPETVETVTTTAQETKGR-----IOTKEVAIKTT 287
Db 174 LAALNLEERKRLVAQFLAESSOR-----SQEAASQREPSADPVIKSKTSQKYMAL 223
Qy 288 LKELVHKRVTSPEWMMQPDYSIENMAOPGGENLKNLTLEICTLTLARTKTAFLDILEK 347
Db 224 VNWLVHGHTSEKQWIQENQESYLSFNSTGNSRSQIKAAALDNATKIMSLTKSAVDY---- 279
Qy 348 AETSKLNTSLPD-----TRTCRIFAFHGMN--YV-KVCHAI CCVLRQGGKRNVTLPFGP 400
Db 280 -----LVGSSVPEIDISKRIWQIFEMNGYDPAVAGSILYGCW---QRSFNKRNVTWLYGP 331
Qy 401 ASTGKSIIAQAATAQAGVNGVCYNAANVFPNDCTNKNLIWEEAGNFGQGVNQFKAICS 460
Db 332 ATTGKTNIAEAIHTVFFYGCNVWNTNFFNDCVDMKLIWEEGKMTNKVVSAAKAILG 391
Qy 461 GQIRIDQKQKSGKQIEPTPVIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPL 520
Db 392 GSKVRVDQKCKSSVQIDSTPVIIVTSNTNMCVVVDGNSSTTFEHQHPLEDRMFKELTKRLP 451
Qy 521 GDFGLVDKNWPMICAWLVKNGYQSTWASYCAKWKGVDPWSE-----NWAEPKVP 570
Db 452 PDFGKITQKQEVKDFFAW-----AKVNOVPVTHEFKVPRELATGKAESKSLK 497
```

Db 452 PDGKITKQEVKDFFAW-----AKNQVPVTHFKVPRELKAGTKGAEKSLK 497

QY 571 TPI-----NLLGSARSPT--TPKSTPLSQNYA-LTPL 600

Db 498 RPLGVDVNTSYKSLKRAKLSFVFPETPRSSDVTVDPAPLRPL 539

RESULT 6

US-10-023-208-21

; Sequence 21, Application US/10023208

; Publication No. US20030124537A1

; GENERAL INFORMATION:

; APPLICANT: Li, Min

; APPLICANT: Liu, Yuan-Ching

; TITLE OF INVENTION: PROCARYOTIC LIBRARIES AND USES

; FILE REFERENCE: A-70174-1/RFT/RMS/RMK

; CURRENT APPLICATION NUMBER: US/10/023,208

; CURRENT FILING DATE: 2001-12-17

; PRIOR APPLICATION NUMBER: US 60/256,163

; PRIOR FILING DATE: 2000-12-14

; NUMBER OF SEQ ID NOS: 63

; SOFTWARE: Patent version 3.1

; SEQ ID NO 21

; LENGTH: 610

; TYPE: PRT

; ORGANISM: adeno-associated virus 5

US-10-023-208-21

Query Match 12.5%; Score 455.5; DB 14; Length 610;

Best Local Similarity 27.8%; Pred. No. 1.2e-30;

Matches 162; Conservative 78; Mismatches 217; Indels 125; Gaps 23;

QY 67 WQSEDMEWETTVDKTKQVFIFDLSLVKKCLFEVLNTKNIFFGDVNVFVQHWGKDQGW 126

Db 35 WELPPESDLNLTLVE--QPQLTVADRIRRVFLYE---WNKFSKQESKFFVQFEKGSY-F 88

QY 127 HCHVLIGKDFP-SQAQGWRRQLNVYNSRWLVTACNVOLTPAERIKLREI-AEDNEWT 184

Db 89 HLHVLVETSGISSMWLGRY-----VSQIRAOVL---KVVFQIEPQINDWA 132

QY 185 LUTYKHQKDYTKCVLFGNMIAVYFLTKKISTSPPRDGGYFLSSDSGWK-TNF- 239

Db 133 I-----TKVKGANKVDGSIYPAVLLPK-----VQPELOWANTNLDYK 173

QY 240 -----LKEGERHLVSKLYTDDMRPETVETVTTAQTGR-----IQTKEVAIKTT 287

Db 174 LAALNLEERKRLVAQFLAESSQR-----SQEASQRBFSADPVIKSKTSQKYMAL 223

QY 288 LKELVHKRVTSPPDWMMPQDSYIEMMAOPGGENLLKNTLEICTLTARTKTAFDLILEK 347

Db 224 VNWLVHGTSEKQWIQENQESYLSFNSTGNSRSQIKAAALDNATKIMSLTKSAVDY- 279

QY 348 AETSKLTNFSLPD---TTCRIFAFHGWN--YV-KVCHAI CCVLNRQGGKRTNVLPHGP 400

Db 280 -----LVGSSVPEDISKRIWQIFEMNGYDPAVAGSILYGC--QRSFNKRTNVLWLYGP 331

QY 401 ASTGKSIIAQAIAQAVNGVCYNAANVPFPNDCTNKNLIWVEEAGNFQGVQVQFKAICS 460

Db 332 ATTGKTINAEIAHTVPFYGCYVNTNENFPFNDVCKMLIWEEGKNTKVVESAKAILG 391

QY 461 GQIRIDQKGSQKLEPTPVMTNENITVVRIGCEERPEHTQPTDRMLNHLTHLPL 520

Db 392 GSKVRVDQKSSVQIDSTPVI VTSNTNMCVVVDGNSSTTFEHOQPLEDRMFKFELTKRLP 451

QY 521 GDFGLVDKNWPMICAWLVKNGYQSTMASYCAKWKVPDQWSE-----NWABPKVP 570

Db 452 PDGKITKQEVKDFFAW-----AKNQVPVTHFKVPRELKAGTKGAEKSLK 497

QY 571 TPI-----NLLGSARSPT--TPKSTPLSQNYA-LTPL 600

Db 498 RPLGVDVNTSYKSLKRAKLSFVFPETPRSSDVTVDPAPLRPL 539

RESULT 7

US-10-375-192-7

; Sequence 7, Application US/10375192

; Publication No. US20030224404A1

; GENERAL INFORMATION:

; APPLICANT: Vega, Manuel

; APPLICANT: Dittanti, Lila

; TITLE OF INVENTION: HIGH-THROUGHPUT DIRECTED EVOLUTION OF NUCLEIC ACIDS BY RATIONAL

; TITLE OF INVENTION: MUTAGENESIS

; FILE REFERENCE: 37851-918

; CURRENT APPLICATION NUMBER: US/10/375,192

; CURRENT FILING DATE: 2003-02-24

; PRIOR APPLICATION NUMBER: US 60/360,085

; PRIOR FILING DATE: 2002-02-25

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7

; LENGTH: 610

; TYPE: PRT

; ORGANISM: Adeno-associated virus 5

US-10-375-192-7

Query Match 12.5%; Score 455.5; DB 14; Length 610;

Best Local Similarity 27.8%; Pred. No. 1.2e-30;

Matches 162; Conservative 78; Mismatches 217; Indels 125; Gaps 23;

QY 67 WQSEDMEWETTVDKTKQVFIFDLSLVKKCLFEVLNTKNIFFGDVNVFVQHWGKDQGW 126

Db 35 WELPPESDLNLTLVE--QPQLTVADRIRRVFLYE---WNKFSKQESKFFVQFEKGSY-F 88

QY 127 HCHVLIGKDFP-SQAQGWRRQLNVYNSRWLVTACNVOLTPAERIKLREI-AEDNEWT 184

Db 89 HLHVLVETSGISSMWLGRY-----VSQIRAOVL---KVVFQIEPQINDWA 132

QY 185 LUTYKHQKDYTKCVLFGNMIAVYFLTKKISTSPPRDGGYFLSSDSGWK-TNF- 239

Db 133 I-----TKVKGANKVDGSIYPAVLLPK-----VQPELOWANTNLDYK 173

QY 240 -----LKEGERHLVSKLYTDDMRPETVETVTTAQTGR-----IQTKEVAIKTT 287

Db 174 LAALNLEERKRLVAQFLAESSQR-----SQEASQRBFSADPVIKSKTSQKYMAL 223

QY 288 LKELVHKRVTSPPDWMMPQDSYIEMMAOPGGENLLKNTLEICTLTARTKTAFDLILEK 347

Db 224 VNWLVHGTSEKQWIQENQESYLSFNSTGNSRSQIKAAALDNATKIMSLTKSAVDY- 279

QY 348 AETSKLTNFSLPD---TTCRIFAFHGWN--YV-KVCHAI CCVLNRQGGKRTNVLPHGP 400

Db 280 -----LVGSSVPEDISKRIWQIFEMNGYDPAVAGSILYGC--QRSFNKRTNVLWLYGP 331

QY 401 ASTGKSIIAQAIAQAVNGVCYNAANVPFPNDCTNKNLIWVEEAGNFQGVQVQFKAICS 460

Db 332 ATTGKTINAEIAHTVPFYGCYVNTNENFPFNDVCKMLIWEEGKNTKVVESAKAILG 391

QY 461 GQIRIDQKGSQKLEPTPVMTNENITVVRIGCEERPEHTQPTDRMLNHLTHLPL 520

Db 392 GSKVRVDQKSSVQIDSTPVI VTSNTNMCVVVDGNSSTTFEHOQPLEDRMFKFELTKRLP 451

QY 521 GDFGLVDKNWPMICAWLVKNGYQSTMASYCAKWKVPDQWSE-----NWABPKVP 570

Db 452 PDGKITKQEVKDFFAW-----AKNQVPVTHFKVPRELKAGTKGAEKSLK 497

QY 571 TPI-----NLLGSARSPT--TPKSTPLSQNYA-LTPL 600

Db 498 RPLGVDVNTSYKSLKRAKLSFVFPETPRSSDVTVDPAPLRPL 539

RESULT 8

US-09-792-630-29

; Sequence 29, Application US/09792630

; Patent No. US20020168640A1

; GENERAL INFORMATION:

; APPLICANT: Li, Min


```
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Muscovy duck parvovirus
US-10-080-376-29

Query Match      12.4%; Score 452.5; DB 13; Length 626;
Best Local Similarity 25.2%; Pred. No. 2.3e-30;
Matches 164; Conservative 81; Mismatches 265; Indels 141; Gaps 22;

QY 75 WETT-----VDENKKQVFIFDSLVKCLFEVLNTKNIF-----PGDVMNFVQHEWGK 122
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 45 WEPTGIMNHEVNLPMVTLADKI-----KNIFIQRNQFNQDSTDFFFQLEEGS 93
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 123 DQGWCHVLIGKQFSQAQGWRRQLNVYRWLVTACNVQLTPAEIK---LREIABD 179
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 94 EY-IHLHAVCPGECRSFVLGRY-----MSQIKDSILRDVYBG 129
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 180 N-----EWVTLTYKHQTKDYTKCVLFGNMIAYVFLTKKISTSPPRDGGVFLSSDSG 234
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 130 KQVKIPDWSITTKRGGQNKVTAA-----YILHYLIPKKQ-----PELOWAFTNMPLF 179
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 235 WKTNFKEGERHLVSKLYTDDMRPETVETVTTTAQETKRGRIQTKKEVAIKTKLKELVHK 294
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 180 TAAALCLOKQELLDAFOSEMNNAVQEDQASTAAPL-----ISNRAAKYNSNLVDWLIEM 235
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 295 RVTSPEDNMMOPDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFTDLILEKAETSKLT 354
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 236 GITSEKQMLTENKESYRSFOATSSNNRQVKAALENARAEMLLTKTATDYLIKGDPVLDIT 295
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 355 NPSLPDTRTCRIFAFHGMVYVCHAI CCVLNRQGGKRNVLPHGPASTGKSIIAQAIQ 414
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 296 K-----NRIYQILKLNINPQYVGVLCGWVXREFNKRNAIWLGPATTKGTNIAEIAH 350
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 415 AVGNVCYNAANVPFNDCTNKNLIWVEEAGNFGQOVNQFKAICSGQTIRIDQKGGSK 474
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 351 AVPFYGVVNTNENFFNDQVDMLIWEEGKMTNKVESAAILGGSVAVRVDQCKGVS 410
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 475 QIEPTPVIMTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDGLVDKNE-----530
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 411 CIEPTVITNTDMCMIVDGNSTTMEHRIPLERMFQIVLSHKGLENGFKISKKEVKPF 470
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 531 --W-----PMICAWLVKNGYSTMASYCAKWKGVDPWSEN-----WAEPK-----568
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 471 FKWANDNLVPVSEFKVPTNEQTKLTE-----PVPERANESEPPKIWAPPTRELEEI 524
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 569 -----VPTPINLLGSARSPTTPKSTPLSONYALTPLASDLEDLALEPWPSTPNT 617
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 525 LRASPELFAVAFLP-----SSPDTSPKRTKTRGEYQVRCAMHSL-DNSMNVFECLEC 576
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 618 PVAGTAETQNTGEAGSKAC-QDQGLSPWTWSEIEEDLRACFAGBPLKKDFSE 667
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 577 ERANPFQFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 612
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

RESULT 11
US-10-082-671-35
; Sequence 35, Application US/10082671
; Publication No. US20030049647A1
; GENERAL INFORMATION:
; APPLICANT: DAHIVAT, BASSIL
; APPLICANT: LI, MIN
; TITLE OF INVENTION: USE OF NUCLEIC ACID LIBRARIES TO CREATE TOXICOLOGICAL
; FILE REFERENCE: XEN/001
; CURRENT APPLICATION NUMBER: US/10/082, 671
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/270,781
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 58
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Muscovy duck parvovirus
US-10-082-671-35

Query Match      12.4%; Score 452.5; DB 14; Length 626;
Best Local Similarity 25.2%; Pred. No. 2.3e-30;
Matches 164; Conservative 81; Mismatches 265; Indels 141; Gaps 22;

QY 75 WETT-----VDENKKQVFIFDSLVKCLFEVLNTKNIF-----PGDVMNFVQHEWGK 122
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 45 WEPTGIMNHEVNLPMVTLADKI-----KNIFIQRNQFNQDSTDFFFQLEEGS 93
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 123 DQGWCHVLIGKQFSQAQGWRRQLNVYRWLVTACNVQLTPAEIK---LREIABD 179
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 94 EY-IHLHAVCPGECRSFVLGRY-----MSQIKDSILRDVYBG 129
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 180 N-----EWVTLTYKHQTKDYTKCVLFGNMIAYVFLTKKISTSPPRDGGVFLSSDSG 234
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 130 KQVKIPDWSITTKRGGQNKVTAA-----YILHYLIPKKQ-----PELOWAFTNMPLF 179
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 235 WKTNFKEGERHLVSKLYTDDMRPETVETVTTTAQETKRGRIQTKKEVAIKTKLKELVHK 294
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 180 TAAALCLOKQELLDAFOSEMNNAVQEDQASTAAPL-----ISNRAAKYNSNLVDWLIEM 235
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 295 RVTSPEDNMMOPDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFTDLILEKAETSKLT 354
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 236 GITSEKQMLTENKESYRSFOATSSNNRQVKAALENARAEMLLTKTATDYLIKGDPVLDIT 295
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 355 NPSLPDTRTCRIFAFHGMVYVCHAI CCVLNRQGGKRNVLPHGPASTGKSIIAQAIQ 414
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 296 K-----NRIYQILKLNINPQYVGVLCGWVXREFNKRNAIWLGPATTKGTNIAEIAH 350
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 415 AVGNVCYNAANVPFNDCTNKNLIWVEEAGNFGQOVNQFKAICSGQTIRIDQKGGSK 474
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 351 AVPFYGVVNTNENFFNDQVDMLIWEEGKMTNKVESAAILGGSVAVRVDQCKGVS 410
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 475 QIEPTPVIMTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDGLVDKNE-----530
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 411 CIEPTVITNTDMCMIVDGNSTTMEHRIPLERMFQIVLSHKGLENGFKISKKEVKPF 470
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 531 --W-----PMICAWLVKNGYSTMASYCAKWKGVDPWSEN-----WAEPK-----568
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 471 FKWANDNLVPVSEFKVPTNEQTKLTE-----PVPERANESEPPKIWAPPTRELEEI 524
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 569 -----VPTPINLLGSARSPTTPKSTPLSONYALTPLASDLEDLALEPWPSTPNT 617
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 525 LRASPELFAVAFLP-----SSPDTSPKRTKTRGEYQVRCAMHSL-DNSMNVFECLEC 576
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 618 PVAGTAETQNTGEAGSKAC-QDQGLSPWTWSEIEEDLRACFAGBPLKKDFSE 667
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 577 ERANPFQFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 612
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

RESULT 12
US-10-097-100-29
; Sequence 29, Application US/10097100
; Publication No. US20030068649A1
; GENERAL INFORMATION:
; APPLICANT: LI, Min
; APPLICANT: Melander, Christian
; APPLICANT: Liu, Hong-Xiang
; APPLICANT: Jin, Cheng He
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FUSION L
; FILE REFERENCE: A-70814/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/097,100
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/953,351
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/232,960
; PRIOR FILING DATE: 2000-09-14
```

; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 29
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Muscovy duck parvovirus
US-10-097-100-29

Query Match 12.4%; Score 452.5; DB 14; Length 626;
Best Local Similarity 25.2%; Pred. No. 2.3e-30;
Matches 164; Conservative 81; Mismatches 265; Indels 141; Gaps 22;

QY 75 WETT-----VDEMTKKQVFIFDSLVKKCLFEVLNTKNIF-----PGDVNFWVQHEWGK 122
DB 45 WEPTGIWNEHVNLPMTLADKI-----KNIFQRMNQFNQDETDFFFOLEEGS 93
QY 123 DOGWCHVILGGDFSOAQGKWRRLQNVVWSRWLTACNVQLTPAERIK---LREIAED 179
DB 94 EY-IHLHVCPCGECRSFVLGRY-----MSQIKDSILRDVYEG 129
QY 180 N-----EWTLTYKHQKQDYTKCVLFGNMIAYVFLTKKISTSPRRDGGYFLSSDSG 234
DB 94 EY-IHLHVCPCGECRSFVLGRY-----MSQIKDSILRDVYEG 129
QY 130 KQVKIPDWFISITTKRGGQNKVTAA-----YILHYLIPKKQ-----PELQWAFNTNPLF 179
DB 235 WKTNFKGGERHLVSKLYTDDMRPETVETVTTAQTGRIQTKKEVAIKTKLKVHK 294
DB 180 TAAALCLOKQRELLDAFQSEMNNAVQEDQASTAAPL-----ISNRAAKNSNLVDWLIEM 235
QY 295 RVTSPEDMMQPDSDYIENMAOPGGENLLKNTLEICTLTARTKTAFDILILEKAETSKLT 354
DB 236 GITSEKQWLITENKESYRSFQATSSNNRQVKAALENARAEMLLTKTATDYLIGKDPVLDT 295
QY 355 NPSLPDTRTCRIFAFHGMNVYVCHACCVLNRQGGKRNVTLPFHGPASTGKSIIAOAIQ 414
DB 296 K-----NRIYQILKLNYPQVYVGLCGWVKEFNKRNAILWYGPAITGKTINIAEIAH 350
QY 415 AVGNVCYNAANVPFNDCTNKLIVBEAGNFGQVNOFKAICSGQTIRIDQKKGSK 474
DB 351 AVPPYGVNWTNENFPFNDCTNKLIVBEAGNFGQVNOFKAICSGQTIRIDQKKGSK 410
QY 475 QIEPTVIMTNTENITVVRIGCEERPEHTQPIRDRMLNTHLTHLPGDFGLVDKNE--- 530
DB 411 CIEPTVITNTDNCMIVDGNSSTTMEHRIPLEERFQIVLSHKLGNFGKISKKEVKEF 470
QY 531 --W-----PMICAMLVKNGYOSTMASYCAKMGKVPDWSN-----WAEPK----- 568
DB 471 FKWANDNLVPVSEFKVPTNEQTKLTE-----PVPERANEPSEPKIWPPTREELEBI 524
QY 569 -----VPTPINLIGSARSPTTPKSTPLSQNYALTPLASDLLELALPSTPNT 617
DB 525 LRASPELFASVAPLP-----SSPDTSPKRRKTRGEYQVRCAMHSL-DNSMNVFECLEC 576
QY 618 PVAGTAETQNTGEAGSKAC-QDQQLSPTWSEIEEDLRACFGAEPLKKDFSE 667
DB 577 ERANFPFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 612

RESULT 13

US-10-023-208-29
; Sequence 29, Application US/10023208
; Publication No. US20030124537A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Liu, Yuan-Ching
; TITLE OF INVENTION: PROCAROTIC LIBRARIES AND USES
; FILE REFERENCE: A-70174-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/023,208
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/256,163
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 29

; LENGTH: 626
; TYPE: PRT
; ORGANISM: Muscovy duck parvovirus
US-10-023-208-29

Query Match 12.4%; Score 452.5; DB 14; Length 626;
Best Local Similarity 25.2%; Pred. No. 2.3e-30;
Matches 164; Conservative 81; Mismatches 265; Indels 141; Gaps 22;

QY 75 WETT-----VDEMTKKQVFIFDSLVKKCLFEVLNTKNIF-----PGDVNFWVQHEWGK 122
DB 45 WEPTGIWNEHVNLPMTLADKI-----KNIFQRMNQFNQDETDFFFOLEEGS 93
QY 123 DOGWCHVILGGDFSOAQGKWRRLQNVVWSRWLTACNVQLTPAERIK---LREIAED 179
DB 94 EY-IHLHVCPCGECRSFVLGRY-----MSQIKDSILRDVYEG 129
QY 180 N-----EWTLTYKHQKQDYTKCVLFGNMIAYVFLTKKISTSPRRDGGYFLSSDSG 234
DB 130 KQVKIPDWFISITTKRGGQNKVTAA-----YILHYLIPKKQ-----PELQWAFNTNPLF 179
QY 235 WKTNFKGGERHLVSKLYTDDMRPETVETVTTAQTGRIQTKKEVAIKTKLKVHK 294
DB 180 TAAALCLOKQRELLDAFQSEMNNAVQEDQASTAAPL-----ISNRAAKNSNLVDWLIEM 235
QY 295 RVTSPEDMMQPDSDYIENMAOPGGENLLKNTLEICTLTARTKTAFDILILEKAETSKLT 354
DB 236 GITSEKQWLITENKESYRSFQATSSNNRQVKAALENARAEMLLTKTATDYLIGKDPVLDT 295
QY 355 NPSLPDTRTCRIFAFHGMNVYVCHACCVLNRQGGKRNVTLPFHGPASTGKSIIAOAIQ 414
DB 296 K-----NRIYQILKLNYPQVYVGLCGWVKEFNKRNAILWYGPAITGKTINIAEIAH 350
QY 415 AVGNVCYNAANVPFNDCTNKLIVBEAGNFGQVNOFKAICSGQTIRIDQKKGSK 474
DB 351 AVPPYGVNWTNENFPFNDCTNKLIVBEAGNFGQVNOFKAICSGQTIRIDQKKGSK 410
QY 475 QIEPTVIMTNTENITVVRIGCEERPEHTQPIRDRMLNTHLTHLPGDFGLVDKNE--- 530
DB 411 CIEPTVITNTDNCMIVDGNSSTTMEHRIPLEERFQIVLSHKLGNFGKISKKEVKEF 470
QY 531 --W-----PMICAMLVKNGYOSTMASYCAKMGKVPDWSN-----WAEPK----- 568
DB 471 FKWANDNLVPVSEFKVPTNEQTKLTE-----PVPERANEPSEPKIWPPTREELEBI 524
QY 569 -----VPTPINLIGSARSPTTPKSTPLSQNYALTPLASDLLELALPSTPNT 617
DB 525 LRASPELFASVAPLP-----SSPDTSPKRRKTRGEYQVRCAMHSL-DNSMNVFECLEC 576
QY 618 PVAGTAETQNTGEAGSKAC-QDQQLSPTWSEIEEDLRACFGAEPLKKDFSE 667
DB 577 ERANFPFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 612

RESULT 14

US-09-792-630-25
; Sequence 25, Application US/09792630
; Patent No. US20020168640A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Dahlyat, Basil I.
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES
; FILE REFERENCE: A-70295/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/792,630
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 25
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Barbarie duck parvovirus
US-09-792-630-25

Query Match 12.4%; Score 452; DB 9; Length 627;
Best Local Similarity 25.8%; Pred. No. 2.6e-30;
Matches 168; Conservative 77; Mismatches 266; Indels 140; Gaps 23;

QY 75 WETT-----VDEMTKKQVFIPIFDSLVKCLFEVLNTKNIP-----PGDWNWVQHEWKG 122
DB 45 WEPTGIWNHVNLPVMTLADKI-----KNIFIQWVNFQNDQDETFQLEEGS 93

QY 123 DQGWCHVLIGGKDFSAQGWRRQLNVYWSRWLTACNVQLTPAERIK---LREIAD 179
DB 94 EY-IHLHCCI-----AQNVRSFVLGRYMS-----QIKDSILRDVYEG 130

QY 180 N-----EWTLLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPPRDGGYFLSSDSG 234
DB 131 KQVKIPDWFISITTKRGQNKVTAA-----YILHYLIPKQ-----PELOWAFTNMPLE 180

QY 235 WKTNFKGHRHLVSKLYTDDMRPETVETVTTAQTCKGRIGRIQTKKEVAIKTTLKELVHK 294
DB 181 TAAALCLQKRELDAFOSENNVAVQEDQASTAAPL-----ISNRAAKNYSNLVDWLIEM 236

QY 295 RVTSPEMMQPDSDYIEMMAQPGGENLLKNTLEICTLTARTKTAFTDLILEKAETSKLT 354
DB 237 GIITSEKQWLTENKESYRSFOATSSNNRQVKAALENARAEMLLTKTATDYLGKDPVLDIT 296

QY 355 NFSLPDTRTCRIFAFHGMNYYKVCHAI CCVLNRQGGKRNVTFLPHGPASTGKSIIAQAIQ 414
DB 297 K-----NRIYQILKLNYPQYVGVSLCGWVKREFNKNALWLYGPATTGKTNI AEIAH 351

QY 415 AVGNVCYNAANVPFNDCNTKNLIWEEAGNFGQOVNOFKAI CSQTIRIDOKGKSGK 474
DB 352 AVPFYGCNVNTNENPFNDCVDKMLIWEEGKNTKVVESAKAILGGSVAVRDCKGKSV 411

QY 475 QIEPTPVIMTNTNITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNE---- 530
DB 412 CIEPTPVITSDMCMIVDGNSTTWEHRIPLEERMFQIVLSHKLGNFGKISKKEVKEF 471

QY 531 --W-----PMICAWLVKNGYQSTWASYCAKWKGVDPDWSN-----WAEPK----- 568
DB 472 FKWANDNLVPVSEFKVPTNEQTKLE-----PVPERANEPSEPPKIWAPPTREELEEI 525

QY 569 -----VPTPINLLGSARSPTTPKSTPLSQNYALTPLASDLEDLALEPWSPTNT 617
DB 526 LRASPELFASVAPLP-----SSPDTSPKRRKTRGEYQVRCAMHSL-DNSMNVFECLEC 577

QY 618 PVAGTAETQNTGEAGSKAC-QDQQLSPTWSETEEDLRACFGAEPLKDPSE 667
DB 578 ERANFPFQSLGE---NFCNQH-----WYD-----CAFCNELKDDMNE 613

RESULT 15
US-09-953-351-25
; Sequence 25, Application US/09953351
; Publication No. US20030036643A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Melander, Christian
; APPLICANT: Liu, Hong-Xiang
; APPLICANT: Jin, Cheng He
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FUSION I
; FILE REFERENCE: A-70814/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/953,351
; CURRENT FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/232,960
; PRIOR FILING DATE: 2000-09-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 25
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Barbarie duck parvovirus
US-09-953-351-25

Query Match 12.4%; Score 452; DB 10; Length 627;

Best Local Similarity 25.8%; Pred. No. 2.6e-30;
Matches 168; Conservative 77; Mismatches 266; Indels 140; Gaps 23;

QY 75 WETT-----VDEMTKKQVFIPIFDSLVKCLFEVLNTKNIP-----PGDWNWVQHEWKG 122
DB 45 WEPTGIWNHVNLPVMTLADKI-----KNIFIQWVNFQNDQDETFQLEEGS 93

QY 123 DQGWCHVLIGGKDFSAQGWRRQLNVYWSRWLTACNVQLTPAERIK---LREIAD 179
DB 94 EY-IHLHCCI-----AQNVRSFVLGRYMS-----QIKDSILRDVYEG 130

QY 180 N-----EWTLLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPPRDGGYFLSSDSG 234
DB 131 KQVKIPDWFISITTKRGQNKVTAA-----YILHYLIPKQ-----PELOWAFTNMPLE 180

QY 235 WKTNFKGHRHLVSKLYTDDMRPETVETVTTAQTCKGRIGRIQTKKEVAIKTTLKELVHK 294
DB 181 TAAALCLQKRELDAFOSENNVAVQEDQASTAAPL-----ISNRAAKNYSNLVDWLIEM 236

QY 295 RVTSPEMMQPDSDYIEMMAQPGGENLLKNTLEICTLTARTKTAFTDLILEKAETSKLT 354
DB 237 GIITSEKQWLTENKESYRSFOATSSNNRQVKAALENARAEMLLTKTATDYLGKDPVLDIT 296

QY 355 NFSLPDTRTCRIFAFHGMNYYKVCHAI CCVLNRQGGKRNVTFLPHGPASTGKSIIAQAIQ 414
DB 297 K-----NRIYQILKLNYPQYVGVSLCGWVKREFNKNALWLYGPATTGKTNI AEIAH 351

QY 415 AVGNVCYNAANVPFNDCNTKNLIWEEAGNFGQOVNOFKAI CSQTIRIDOKGKSGK 474
DB 352 AVPFYGCNVNTNENPFNDCVDKMLIWEEGKNTKVVESAKAILGGSVAVRDCKGKSV 411

QY 475 QIEPTPVIMTNTNITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNE---- 530
DB 412 CIEPTPVITSDMCMIVDGNSTTWEHRIPLEERMFQIVLSHKLGNFGKISKKEVKEF 471

QY 531 --W-----PMICAWLVKNGYQSTWASYCAKWKGVDPDWSN-----WAEPK----- 568
DB 472 FKWANDNLVPVSEFKVPTNEQTKLE-----PVPERANEPSEPPKIWAPPTREELEEI 525

QY 569 -----VPTPINLLGSARSPTTPKSTPLSQNYALTPLASDLEDLALEPWSPTNT 617
DB 526 LRASPELFASVAPLP-----SSPDTSPKRRKTRGEYQVRCAMHSL-DNSMNVFECLEC 577

QY 618 PVAGTAETQNTGEAGSKAC-QDQQLSPTWSETEEDLRACFGAEPLKDPSE 667
DB 578 ERANFPFQSLGE---NFCNQH-----WYD-----CAFCNELKDDMNE 613

Search completed: January 22, 2005, 04:03:40
Job time : 93.6667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:27:18 ; Search time 22.3333 Seconds
(without alignments)
2895.119 Million cell updates/sec

Title: US-10-069-056-6

Perfect score: 3638

Sequence: 1 MAGNAYSDEVLGATNWLKXK.....RACFGAEPLKQFSEPLNLD 672

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB ID	Description
1	3625	99.6	672	1	UYPVIM	noncapsid protein
2	3558	97.8	721	1	UYPVIM	noncapsid protein
3	3328	91.5	672	1	UYPVIM	noncapsid protein
4	3324	91.4	668	1	UYPVME	noncapsid protein
5	2649	72.8	668	1	UYPVME	noncapsid protein
6	2646	72.7	668	1	UYPVFP	noncapsid protein
7	2644	72.7	668	1	UYPVCP	noncapsid protein
8	2487	68.4	660	1	UYPVPP	noncapsid protein
9	2478	68.1	662	1	UYPVNA	noncapsid protein
10	1623.5	44.6	392	1	UYPVIF	noncapsid protein
11	550	15.1	641	2	S41439	gene NS-1 protein
12	547.5	15.0	641	2	S41434	gene NS-1 protein
13	543	14.9	620	1	UYPVAP	noncapsid protein
14	533.5	14.7	641	2	S41861	gene NS-1 protein
15	452.5	12.4	626	2	S52209	noncapsid protein
16	424.5	11.7	536	1	UYADIA	noncapsid protein
17	351	9.6	726	1	UYPVSI	noncapsid protein
18	347	9.5	671	1	UYPV19	noncapsid protein
19	168	4.6	490	2	T44050	hypothetical prote
20	165.5	4.5	490	2	J01630	noncapsid protein
21	144	4.0	276	2	S26428	hypothetical 31.5K
22	131.5	3.6	849	1	UYPVAD	noncapsid protein
23	125	3.4	545	2	B44054	orf2 protein - Jun
24	124.5	3.4	614	1	W1L41	E1 protein - human
25	122.5	3.4	647	1	W1L339	E1 protein - human
26	120.5	3.3	825	2	G96865	protein F22C12.12
27	120.5	3.3	1008	2	T18832	probable RNA helic
28	114.5	3.1	497	2	C70454	transcription regu
29	113.5	3.1	1098	2	B70232	hypothetical prote

ALIGNMENTS

RESULT 1

UYPVIM

noncapsid protein NS1 - minute virus of mice

C;Species: minute virus of mice, murine parvovirus

C;Date: 14-Nov-1983 #sequence_revision 28-Aug-1985 #text_change 09-Jul-2004

C;Accession: A03696

R;Astell, C.R.; Thomson, M.; Merchlinsky, M.; Ward, D.C.

Nucleic Acids Res. 11, 999-1018, 1983

A;Title: The complete DNA sequence of minute virus of mice, an autonomous parvovirus.

A;Reference number: A03696; MUID:83143341; PMID:6298737

A;Accession: A03696

A;Molecule type: DNA

A;Residues: 1-672 <AST>

A;Cross-references: UNIPROT:P03134; EMBL:V01115

A;Superfamily: parvovirus noncapsid protein

C;Keywords: noncapsid protein

Query Match 99.6%; Score 3625; DB 1; Length 672;

Best Local Similarity 99.7%; Pred. No. 1.5e-255;

Matches 670; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy	1	MAGNAYSDEVLGATNWLKESNQEVFSFVKNNENVOLNGKDIGNWSYKKELODELKSLQ	60
Db	1	MAGNAYSDEVLGATNWLKESNQEVFSFVKNNENVOLNGKDIGNWSYKKELODELKSLQ	60
Qy	61	RGAEITTDQSEDMEWETTVDENTKKQVFFDLSLVKKCLFEVLNTKNIIFPGDVANFVQHEW	120
Db	61	RGAEITTDQSEDMEWETTVDENTKKQVFFDLSLVKKCLFEVLNTKNIIFPGDVANFVQHEW	120
Qy	121	GKDGQWHCHVLIGGKDFSOAQGKWRRLNLYVWSRWLVACNVQLTPAERIKLRETAEDN	180
Db	121	GKDGQWHCHVLIGGKDFSOAQGKWRRLNLYVWSRWLVACNVQLTPAERIKLRETAEDN	180
Qy	181	EWVTLTYKHQTKDYTKCVLFGNMIAYVFLFKKIKISTPPRDGGYFLSSDSGKTNFL	240
Db	181	EWVTLTYKHQTKDYTKCVLFGNMIAYVFLFKKIKISTPPRDGGYFLSSDSGKTNFL	240
Qy	241	KEGERHLVSKLYTDDMRPETVTTTAQETKKGRITQTKKEVAIKTKLKVHKTSP	300
Db	241	KEGERHLVSKLYTDDMRPETVTTTAQETKKGRITQTKKEVAIKTKLKVHKTSP	300
Qy	301	DMMMPDSDVIENMAOPGGENLLKNTLEICTLTLATKTAFDILILEKAETSKLTNFSLPD	360
Db	301	DMMMPDSDVIENMAOPGGENLLKNTLEICTLTLATKTAFDILILEKAETSKLTNFSLPD	360
Qy	361	TRTCRIFAFHGMNYYVVKVCHACVCLNRQGGKRNVLPHGPASTGKSIIAQAIAQAVGNVG	420
Db	361	TRTCRIFAFHGMNYYVVKVCHACVCLNRQGGKRNVLPHGPASTGKSIIAQAIAQAVGNVG	420
Qy	421	CYNAANVNFNDCTNKNLIWVEEAGNFQGVNQFKAICSGQITRIDQKKGSKQIEPTP	480
Db	421	CYNAANVNFNDCTNKNLIWVEEAGNFQGVNQFKAICSGQITRIDQKKGSKQIEPTP	480

```
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLYK 540
|||||
Db 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLYK 540
|||||
QY 541 NGYQSTMASYCAKWKVPDWSNWAEPKVPPTPINILGSARSPTTTPKSTPLSONYALTPL 600
|||||
Db 541 NGYQSTMASYCAKWKVPDWSNWAEPKVPPTPINILGSARSPTTTPKSTPLSONYALTPL 600
|||||
QY 601 ASDLEDLALPWSPTNTPVAGTAETQNTGEGSKACQDQGLSPTWSEIEDLRACFGAEP 660
|||||
Db 601 ASDLEDLALPWSPTNTPVAGTAETQNTGEGSKACQDQGLSPTWSEIEDLRACFGAEP 660
|||||
QY 661 LKXDFSEPLNLD 672
|||
Db 661 LKXDFSEPLNLD 672
|||

RESULT 2
UYPVIM
noncapsid protein NS1 - minute virus of mice (strain VMV1)
C:Species: minute virus of mice, murine parvovirus
C>Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C:Accession: A23008; A29510
R:Sahli, R.; McMaster, G.K.; Hirt, B.
Nucleic Acids Res. 13, 3617-3633, 1985
A>Title: DNA sequence comparison between two tissue-specific variants of the autonomous
A:Reference number: A23008; MUID:85242059; PMID:3855242
A:Accession: A23008
A:Molecule type: DNA
A:Residues: 1-721 <SAH>
A:Cross-references: UNIPROT:Q84363; EMBL:X02481
R:Atell, C.R.; Gardiner, E.M.; Tattersall, P.
J. Virol. 57, 656-669, 1986
A>Title: DNA sequence of the lymphotropic variant of minute virus of mice, VMV(1), and
A:Reference number: A29510; MUID:86115415; PMID:3502703
A:Accession: A29510
A:Molecule type: DNA
A:Residues: 1-645, 1-647-721 <AST>
A:Cross-references: EMBL:M12032; NID:g332289; PIDN:AAA69566.1; PID:g825477
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match 97.8%; Score 3558; DB 1; Length 721;
Best Local Similarity 97.6%; Pred. No. 1.2e-250;
Matches 656; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVLGATNMLKEKSNQEVFSFVKENNVOLNGKDIGNSYKKELQDELSLQ 60
|||||
Db 50 MAGNAYSDEVLGATNMLKEKSNQEVFSFVKEDVOLNGKDIGNSYKKELQDELSLQ 109
|||||
QY 61 RGAETTWQSEDMEWETTVDEMTEKQVFI FDSLVKKCLPEVLTKNIFPGDVNWFVQHEW 120
|||||
Db 110 RGAETTWQSEDMEWETTVDEMTEKQVFI FDSLVKKCLPEVLTKNIAPADVTWVQHEW 169
|||||
QY 121 GKQDQGHCHVLIGGKDFSOAQGKWRRLQNVYWSRMLVTACNVQLTPAERI KLRLEAEN 180
|||||
Db 170 GKQDQGHCHVLIGGKDFSOAQGKWRRLQNVYWSRMLVTACNVQLTPAERI KLRLEAEN 229
|||||
QY 181 EWTLLTYKHKQTKDYTKCVLFGNMIAYYFLTKKISTSPPRDGGYFLSSDSGKNTFL 240
|||||
Db 230 EWTLLTYKHKQTKDYTKCVLFGNMIAYYFLTKKISTSPPRDGGYFLSSDSGKNTFL 289
|||||
QY 241 KEGERHLVSKLYTDDMRPEVETTVTAQETKRGRIQTKKEVAIKTTLKELVHKRVTSPE 300
|||||
Db 290 KEGERHLVSKLYTDDMRPEVETTVTAQETKRGRIQTKKEVAIKTTLKELVHKRVTSPE 349
|||||
QY 301 DNMWMPDSYIEMMAOPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360
|||||
Db 350 DNMWMPDSYIEMMAOPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 409
|||||
QY 361 TRTCRIFAFHGWNYKVCHAI CCVLNRQGGKNTVLFHGPASTGKSI IIAQIAQAVGNVG 420
|||||
```

```
Db 410 TRTCRIFAFHGWNYKVCHAI CCVLNRQGGKNTVLFHGPASTGKSI IIAQIAQAVGNVG 469
|||||
QY 421 CYNAANVNPFNDCTNKNLIWVEEAGNFQGVNQVQKKAICSGQTIRIDQKKGSKQIEPTP 480
|||||
Db 470 CYNAANVNPFNDCTNKNLIWVEEAGNFQGVNQVQKKAICSGQTIRIDQKKGSKQIEPTP 529
|||||
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLYK 540
|||||
Db 530 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLYK 589
|||||
QY 541 NGYQSTMASYCAKWKVPDWSNWAEPKVPPTPINILGSARSPTTTPKSTPLSONYALTPL 600
|||||
Db 590 NGYQSTMASYCAKWKVPDWSNWAEPKVPPTPINILGSARSPTTTPKSTPLSONYALTPL 649
|||||
QY 601 ASDLEDLALPWSPTNTPVAGTAETQNTGEGSKACQDQGLSPTWSEIEDLRACFGAEP 660
|||||
Db 650 ASDLEDLALPWSPTNTPVAGTAETQNTGEGSKACQDQGLSPTWSEIEDLRACFGAEP 709
|||||
QY 661 LKXDFSEPLNLD 672
|||
Db 710 LKXDFSEPLNLD 721
|||

RESULT 3
UYPVIM
noncapsid protein NS1 - parvovirus Hi
C:Species: parvovirus Hi
A>Note: host Homo sapiens (man)
C>Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 09-Jul-2004
C:Accession: A03695
R:Rhode III, S.L.; Paradiso, P.R.
J. Virol. 45, 173-184, 1983
A>Title: Parvovirus genome: nucleotide sequence of H-1 and mapping of its genes by hybrid
A:Reference number: A03695; MUID:83112183; PMID:6823009
A:Accession: A03695
A:Molecule type: DNA
A:Residues: 1-672 <RHO>
A:Cross-references: UNIPROT:P03133; EMBL:X01457; NID:g60993; PIDN:CAA25689.1; PID:g60994
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match 91.5%; Score 3328; DB 1; Length 672;
Best Local Similarity 91.1%; Pred. No. 5.9e-234;
Matches 612; Conservative 25; Mismatches 35; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVLGATNMLKEKSNQEVFSFVKENNVOLNGKDIGNSYKKELQDELSLQ 60
|||||
Db 1 MAGNAYSDEVLGATNMLKEKSNQEVFSFVKENNVOLNGKDIGNSYKKELQDELSLQ 60
|||||
QY 61 RGAETTWQSEDMEWETTVDEMTEKQVFI FDSLVKKCLPEVLTKNIFPGDVNWFVQHEW 120
|||||
Db 61 RGAETTWQSEDMEWESA VDDMTKKQVFI FDSLVKKCLPEVLSTKNIAPSNVTFVQHEW 120
|||||
QY 121 GKQDQGHCHVLIGGKDFSOAQGKWRRLQNVYWSRMLVTACNVQLTPAERI KLRLEAEN 180
|||||
Db 121 GKDPGWHCHVLIGGKDFSQPGKWRRLQNVYWSRMLVTACNVQLTPAERI KLRLEAEN 180
|||||
QY 181 EWTLLTYKHKQTKDYTKCVLFGNMIAYYFLTKKISTSPPRDGGYFLSSDSGKNTFL 240
|||||
Db 181 EWTLLTYKHKQTKDYTKCVLFGNMIAYYFLSKKICTSPPRDGGYFLSSDSGKNTFL 240
|||||
QY 241 KEGERHLVSKLYTDDMRPEVETTVTAQETKRGRIQTKKEVAIKTTLKELVHKRVTSPE 300
|||||
Db 241 KEGERHLVSKLYTDEMKEPETVTVTAQEAQRGRIQTRREVSIKTTLKELVHKRVTSPE 300
|||||
QY 301 DNMWMPDSYIEMMAOPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360
|||||
Db 301 DNMWMPDSYIEMMAOPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLIANSFMS 360
|||||
QY 361 TRTCRIFAFHGWNYKVCHAI CCVLNRQGGKNTVLFHGPASTGKSI IIAQIAQAVGNVG 420
|||||
Db 361 TRTCRIFAFHGWNYKVCHAI CCVLNRQGGKNTVLFHGPASTGKSI IIAQIAQAVGNVG 420
|||||
```

```
QY 421 CYNAANVFPNDCTNKNLIWVEAGNFGQVNOQFKAI CSQTIRIDQKGSQKQIEPTP 480
DB 421 CYNAANVFPNDCTNKNLIWVEAGNFGQVNOQFKAI CSQTIRIDQKGSQKQIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
DB 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
QY 541 NGYOSTMASYCAKWKVPDWSNWAEPKVPPTPNLLGSARSPTTPKSTPLSQNYALTPL 600
DB 541 NGYOSTMASYCAKWKVPDWSNWAEPKVPPTPNLLGSARSPTTPKSTPLSQNYALTPL 600
QY 601 ASLEDLALPWPSTPNTPVAGTAETONTGEAGSKACODGQOLSPTWSEIBEDLACFCAEP 660
DB 601 ASLEDLALPWPSTPNTPVAGTAETONTGEAGSKACODGQOLSPTWSEIBEDLACFCAEP 660
QY 661 LKXDFSEPLNLD 672
DB 661 LESDFNEELTLD 672

RESULT 4
A44276
noncapsid protein NS1 - parvovirus LuIII
C:Species: parvovirus LuIII
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C:Accession: A44276
R:Diifoot, N.; Chen, K.C.; Bates, R.C.; Lederma, M.
Virology 192, 339-345, 1993
A:Title: The complete nucleotide sequence of parvovirus LuIII and localization of a unique
A:Reference number: A44276; MUID:93297126; PMID:8517025
A:Accession: A44276
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-668 <DIF>
A:Cross-references: UNIPROT:P36311; GB:M81888
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match 91.4%; Score 3324; DB 1; Length 668;
Best Local Similarity 91.4%; Pred. No. 1.1e-233;
Matches 614; Conservative 24; Mismatches 30; Indels 4; Gaps 2;

QY 1 MAGNAYSDEVLGATNMKESKNOEVFSFVKPNENVOLNGKDIGNSYKKELODELKSLQ 60
DB 1 MAGNAYSDEVLGATNMKESKNOEVFSFVKPNENVOLNGKDIGNSYKKELODELKSLQ 60
QY 61 RGAETTDQSEDMEWETTVDENTKKQVFI FDSL VKKCLFEVLNTKNI FPGDVNMFVQHEW 120
DB 61 RGAETTDQSEDMEWESSVDLTKQVFI FDSL VKKCLFEVLNTKNI FPGDVNMFVQHEW 120
QY 121 GKQGWCHVLIIGKDFSOAQGWRRQLNLYWSRWLVATCNVQLTPAERIKLREIAEDN 180
DB 121 GKQGWCHVLIIGKDFSOAQGWRRQLNLYWSRWLVATCNVQLTPAERIKLREIAEDQ 180
QY 181 EWTLLTYKHQTKQDYTKVCFGNMIAYYFLTKKKIISTSPPRDGGYFLSSDSGWKTNFL 240
DB 181 EWTLLTYKHQTKQDYTKVCFGNMIAYYFLTKKKIISTSPPRDGGYFLSSDSGWKTNFL 240
QY 241 KEGERHLVSKLYTDDMRPETVETVTVAQETKGRIOQTKKEVAIKTLLKELVHKRVTSPE 300
DB 241 KEGERHLVSKLYTDDMRPETVETVTVAQETKGRIOQTKKEVAIKTLLKELVHKRVTSPE 300
QY 301 DMMWMPDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLLEKAEYSKLTNFSLPD 360
DB 301 DMMWMPDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLLEKAEYSKLTNFSLPD 360
QY 361 TRTCRIFAFHGWNYKVCHAI CCVLRNQGGRNTVLFHGPASTGCKSI IAQAIAQAVGNVG 420
DB 361 TRTCRIFAFHGWNYKVCHAI CCVLRNQGGRNTVLFHGPASTGCKSI IAQAIAQAVGNVG 420
QY 421 CYNAANVFPNDCTNKNLIWVEAGNFGQVNOQFKAI CSQTIRIDQKGSQKQIEPTP 480
```

```
DB 421 CYNAANVFPNDCTNKNLIWVEAGNFGQVNOQFKAI CSQTIRIDQKGSQKQIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
DB 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
QY 541 NGYOSTMASYCAKWKVPDWSNWAEPKVPPTPNLLGSARSPTTPKSTPLSQNYALTPL 600
DB 541 NGYOSTMASYCAKWKVPDWSNWAEPKVPPTPNLLGSARSPTTPKSTPLSQNYALTPL 600
QY 601 ASLEDLALPWPSTPNTPVAGTAETONTGEAGSKACODGQOLSPTWSEIBEDLACFCAEP 660
DB 601 ASLEDLALPWPSTPNTPVAGTAETONTGEAGSKACODGQOLSPTWSEIBEDLACFCAEP 660
QY 661 LKXDFSEPLNLD 672
DB 657 WKSDEQLPNLD 668

RESULT 5
UYPVME
noncapsid protein NS1 - mink enteritis virus (strain Abashiri)
C:Species: mink enteritis virus, MEV
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: A38350
R:Kariatsumari, T.; Horiuchi, M.; Hama, E.; Yaguchi, K.; Ishiguro, N.; Goto, H.; Shinaga
J. Gen. Virol. 72, 867-875, 1991
A:Title: Construction and nucleotide sequence analysis of an infectious DNA clone of the
A:Reference number: A38350; MUID:91202123; PMID:2016597
A:Accession: A38350
A:Molecule type: DNA
A:Residues: 1-668 <KAR>
A:Cross-references: UNIPROT:P27438; GB:D00765; NID:g222435; PIDN:BAA00662.1; PID:g222436
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match 72.8%; Score 2649; DB 1; Length 668;
Best Local Similarity 73.1%; Pred. No. 1.4e-184;
Matches 491; Conservative 72; Mismatches 103; Indels 6; Gaps 5;

QY 1 MAGNAYSDEVLGATNMKESKNOEVFSFVKPNENVOLNGKDIGNSYKKELODELKSLQ 60
DB 1 MSGNQITEEVMEGVNWLKHAENEAFSVFKCDNVQLNGKDVHNNYTKFIQNEELTSLI 60
QY 61 RGAETTDQOS--EDMEWETTVDENTKKQVFI FDSL VKKCLFEVLNTKNI FPGDVNMFVQH 118
DB 61 RGAETAMDQTEBEEBMDWESEVDSLAKKQVQTFDALIKKCLFEVFSKNIEPNECWFIQH 120
QY 119 EWGKDGWHCHVLIIGKDFSOAQGWRRQLNLYWSRWLVATCNVQLTPAERIKLREIAE 178
DB 121 EWGKDGWHCHVLIIGKDFSOAQGWRRQLNLYWSRWLVATCNVQLTPAERIKLREIAE 180
QY 179 DNEWVTLTYKHQTKQDYTKVCFGNMIAYYFLTKKKIISTSPPRDGGYFLSSDSGWKTN 238
DB 181 DSEWVTLTYRHQTKQDYTKVCFGNMIAYYFLTKKKI--VHMTKESGYFLSTDGSHKFN 239
QY 239 FLKEGERHLVSKLYTDDMRPETVETVTVAQETKGRIOQTKKEVAIKTLLKELVHKRVTS 298
DB 240 FMKYQDRHVTSLTYTSQMKPETVETVTVAQETKGRIOQTKKEVSIKTLRDLVSKRVTS 299
QY 299 PEDMMQPDYSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLLEKAEYSKLTNFSL 358
DB 300 PEDMMQPDYSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLLEKAEYSKLTNFSL 359
QY 359 PDTRTCRIFAFHGWNYKVCHAI CCVLRNQGGRNTVLFHGPASTGCKSI IAQAIAQAVGN 418
DB 360 ANSRTCOIFRMHGMNWKVCHAIACVLRNQGGRNTVLFHGPASTGCKSI IAQAIAQAVGN 419
QY 419 VGCYNAANVFPNDCTNKNLIWVEAGNFGQVNOQFKAI CSQTIRIDQKGSQKQIEP 478
DB 420 VGCYNAANVFPNDCTNKNLIWVEAGNFGQVNOQFKAI CSQTIRIDQKGSQKQIEP 479
QY 479 TPVIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWL 538
```

Db 480 TPVIMTNNITVIRIGCEERPEHTQPIRDRMLNKLKLVCKLPGDFGLVDKEEPLICAWL 539
QY 539 VKNGYOSTWASYCAKWKGVDPWSENWAEKVPPTPINLLGSARSPFTTPKSTPLSQNYALT 598
Db 540 VRHGYESTWANYTHHWGKVPEDENWAEPKIQGVNSPG-CKDLETQAASNPQSDHVL 598
QY 599 PLASDLEDLALPWPSTPNTPVAGTAEONTGEAGSKACODGQLSPTWSEIEDLRACFGA 658
Db 599 PLTPDVVDLALPWPSTPNTPIAETA-NQOSNOLG-VTHKDVQASPTWSEIADLRAIFTS 656
QY 659 EPLKDFSEPLN 670
Db 657 EQLEEDFRDLD 668

RESULT 6

UYPVFP

noncapsid protein NS1 - feline panleukopenia virus (strain 193)

C:Species: feline panleukopenia virus, FPLV

C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004

C:Accession: A36608

R:Martin, J.C.; Davidson, B.E.; Studdert, M.J.

J. Gen. Virol. 71, 2747-2753, 1990

A:Title: Nucleotide sequence of feline panleukopenia virus: comparison with canine parvovirus

A:Reference number: A36608; MUID:91073139; PMID:2174965

A:Accession: A36608

A:Molecule type: DNA

A:Residues: 1-668 <VAR>

A:Cross-references: UNIPROT:P24842; GB:X55115; NID:960863; PIDN:CAA38910.1; PID:960864

C:Superfamily: parvovirus noncapsid protein

C:Keywords: noncapsid protein

Query Match 72.7%; Score 2646; DB 1; Length 668;

Best Local Similarity 73.1%; Pred. No. 2.3e-184;

Matches 491; Conservative 72; Mismatches 103; Indels 6; Gaps 5;

QY 1 MAGNAYSDEVLGATNWLKESNOEVFSFVKFNENVLQNGKDIGNSYKKELQBELKSLQ 60
Db 1 MSGNQITEEVMEGVNWLKHAENAFSFFKCDNVQLNGKDVRRWNYTKPIQNEELTSLI 60
QY 61 RGAETTWDS--EDMEWETTVDMTKKQVIFDLSLVKKCLFEVLNTKNIFFPGDVNFWOH 118
Db 61 RGAQTAMDOTEEEMDWESEVDSLAKKQVOTFDALIKCLFEVFSKNIEPNECVWFIOH 120
QY 119 EWKDGQGHCHVLIGKDFSOAGKWRRLQNVYWSRWLVACNVOLTPAERIKLREIAE 178
Db 121 EWKDGQGHCHVLHSHKLNLOQATGKMLRRQNNYWSRWLVTLCSVNLTPPEKIKLREIAE 180
QY 179 DNEWVTLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPRPGGYFLSSDSGKWN 238
Db 181 DSEWVTLTYRHQTKDYKVMVHFGNMIAYFLTKKI-VHMTKESGYFLSTDGSKFN 239
QY 239 FLKEGERHLVSKLYTDDMRPETVETVTTAQTETKRGRIQTKKEVAIKTTLKELVHGRVTS 298
Db 240 FMKYQDRHTVSTLYTEQMKPETVETVTTAQTETKRGRIQTKKEVSIKTLRDLVSKRVTS 299
QY 299 PEDWMMQPDYSIEMMAQPGENLLKNTLEICTLTARTKTAFLDILEKAEISKLTNPSL 358
Db 300 PEDWMMQPDYSIEMMAQPGENLLKNTLEICTLTARTKTAFLDILEKADNTKLTNFDL 359
QY 359 PDTRTCRIFAFHGNVYKVKCHACVCLNROGKRNVTLPFGPASTGKSIIAQAIAQAVGN 418
Db 360 ANSRTQCFRMHGNWIKVCHACVCLNROGKRNVTLPFGPASTGKSIIAQAIAQAVGN 419
QY 419 VGCYNAANVPFPNDCTNKNLIWVEEAGNFQQVNOQFKAI CSQTTRIDQKGGSKQIEP 478
Db 420 VGCYNAANVPFPNDCTNKNLIWEEAGNFQQVNOQFKAI CSQTTRIDQKGGSKQIEP 479
QY 479 TPVIMTNNITVIRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNWPMICAWL 538
Db 480 TPVIMTNNITVIRIGCEERPEHTQPIRDRMLNKLKLVCKLPGDFGLVDKEEPLICAWL 539

QY 539 VKNGYOSTWASYCAKWKGVDPWSENWAEKVPPTPINLLGSARSPFTTPKSTPLSQNYALT 598
Db 540 VRHGYESTWANYTHHWGKVPEDENWAEPKIQEGI-ISPCKDLETQAASNPQSDHVL 598
QY 599 PLASDLEDLALPWPSTPNTPVAGTAEONTGEAGSKACODGQLSPTWSEIEDLRACFGA 658
Db 599 PLTPDVVDLALPWPSTPNTPIAETA-NQOSNOLG-VTHKDVQASPTWSEIADLRAIFTS 656
QY 659 EPLKDFSEPLN 670
Db 657 EQLEEDFRDLD 668

RESULT 7

UYPVCP

noncapsid protein NS1 - canine parvovirus (strain N)

C:Species: canine parvovirus, CpV

C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004

C:Accession: A29962

R:Reed, A.P.; Jones, E.V.; Miller, T.J.

J. Virol. 62, 266-276, 1988

A:Title: Nucleotide sequence and genome organization of canine parvovirus.

A:Reference number: A29962; MUID:88062992; PMID:2824850

A:Accession: A29962

A:Molecule type: DNA

A:Residues: 1-668 <REE>

A:Cross-references: UNIPROT:P12929; EMBL:M12926; NID:G333438; PIDN:AAA67459.1; PID:G33343

C:Superfamily: parvovirus noncapsid protein

C:Keywords: noncapsid protein

Query Match 72.7%; Score 2644; DB 1; Length 668;

Best Local Similarity 73.1%; Pred. No. 3.3e-184;

Matches 491; Conservative 72; Mismatches 103; Indels 6; Gaps 5;

QY 1 MAGNAYSDEVLGATNWLKESNOEVFSFVKFNENVLQNGKDIGNSYKKELQBELKSLQ 60
Db 1 MSGNQITEEVMEGVNWLKHAENAFSFFKCDNVQLNGKDVRRWNYTKPIQNEELTSLI 60
QY 61 RGAETTWDS--EDMEWETTVDMTKKQVIFDLSLVKKCLFEVLNTKNIFFPGDVNFWOH 118
Db 61 RGAQTAMDOTEEEMDWESEVDSLAKKQVOTFDALIKCLFEVFSKNIEPNECVWFIOH 120
QY 119 EWKDGQGHCHVLIGKDFSOAGKWRRLQNVYWSRWLVACNVOLTPAERIKLREIAE 178
Db 121 EWKDGQGHCHVLHSHKLNLOQATGKMLRRQNNYWSRWLVTLCSVNLTPPEKIKLREIAE 180
QY 179 DNEWVTLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPRPGGYFLSSDSGKWN 238
Db 181 DSEWVTLTYRHQTKDYKVMVHFGNMIAYFLTKKI-VHMTKESGYFLSTDGSKFN 239
QY 239 FLKEGERHLVSKLYTDDMRPETVETVTTAQTETKRGRIQTKKEVAIKTTLKELVHGRVTS 298
Db 240 FMKYQDRQVSTLYTEQMKPETVETVTTAQTETKRGRIQTKKEVSIKTLRDLVSKRVTS 299
QY 299 PEDWMMQPDYSIEMMAQPGENLLKNTLEICTLTARTKTAFLDILEKAEISKLTNPSL 358
Db 300 PEDWMMQPDYSIEMMAQPGENLLKNTLEICTLTARTKTAFLDILEKADNTKLTNFDL 359
QY 359 PDTRTCRIFAFHGNVYKVKCHACVCLNROGKRNVTLPFGPASTGKSIIAQAIAQAVGN 418
Db 360 ANSRTQCFRMHGNWIKVCHACVCLNROGKRNVTLPFGPASTGKSIIAQAIAQAVGN 419
QY 419 VGCYNAANVPFPNDCTNKNLIWVEEAGNFQQVNOQFKAI CSQTTRIDQKGGSKQIEP 478
Db 420 VGCYNAANVPFPNDCTNKNLIWEEAGNFQQVNOQFKAI CSQTTRIDQKGGSKQIEP 479
QY 479 TPVIMTNNITVIRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNWPMICAWL 538
Db 480 TPVIMTNNITVIRIGCEERPEHTQPIRDRMLNKLKLVCKLPGDFGLVDKEEPLICAWL 539
QY 539 VKNGYOSTWASYCAKWKGVDPWSENWAEKVPPTPINLLGSARSPFTTPKSTPLSQNYALT 598
Db 540 VRHGYESTWANYTHHWGKVPEDENWAEPKIQEGINSFG-CKDLETQAASNPQSDQVLT 598

Db 421 YNAANVNPFFNDCTNKNLWTEAGNSQVNOQKACSGQTIKIDQKGSQIEPTPV 480
Qy 482 IMTNNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVKN 541
Db 481 IMTNNEDITKVRIGCEERPEHTQPIRDRMLNINIRKLPDGFGLTEWPLICAWLVKK 540
Qy 542 QYQSTMASYCAKGVKPDWSENWAEKVPPTPINLLGSARSPFTTPKSTPLSQNYALTPLA 601
Db 541 QYQATMASYMHGWNVDWSEKWEKPKQTPINTPDSQIS-TSVKTSPADNNVAAPIQ 599
Qy 602 SDLE-DLALPWSPTNPVAGTASTQNTGEGAGSKACODGQL---SPWSEIIBDLRACFG 657
Db 600 EDLDLALPWSPTNPFTTNLHLTTP-----PDSAIRTPSPWSEIIBDLRACFG 652

RESULT 10

UYPVLF
noncapsid protein NS1 - feline panleukopenia virus (fragment)
C:Species: feline panleukopenia virus, FPLV
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 09-Jul-2004
C:Accession: A03697
R:Carlson, J.; Rushlow, K.; Maxwell, I.; Maxwell, F.; Winston, S.; Hahn, W.
J. Virol. 55, 574-587, 1985
A:Title: Cloning and sequence of DNA encoding structural proteins of the autonomous parvovirus
A:Reference number: A03697; MUID:85265017; PMID:2991581
A:Accession: A03697
A:Molecule type: DNA
A:Residues: 1-392 <CAR>
A:Cross-references: UNIPROT:P06431; EMBL:M10824; NID:g333474; PIDN:AAA47160.1; PID:g333474
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match 44.6%; Score 1623.5; DB 1; Length 392;
Best Local Similarity 77.5%; Pred. No. 2.5e-110;
Matches 306; Conservative 34; Mismatches 52; Indels 3; Gaps 3;

Qy 276 IQTKKEVAIKTKLKHVKTSPEDMMQPDYSIEMMAQPGENLLKNTLBICTLLA 335
Db 1 IQTKKEVSIKTLRLDLSKRVTSPEWMLQPDYSIEMMAQPGENLLKNTLBICTLLA 60
Qy 336 RTKTAFLDLLEKATSKLTNFSLPDTRTCIFAFHGNVYKVCIAACVLRQGGKNTV 395
Db 61 RTKTAFLDLLEKANNTKLTNFDLANSRTQIFRMHGNWIKVCHALACVLRQGGKNTV 120
Qy 396 LFHGPASTGKSIIAQAIAQAVGNVGVNANVFPNDCTNKLIIWVEAGNFGQVNOF 455
Db 121 LFHGPASTGKSIIAQAIAQAVGNVGVNANVFPNDCTNKLIIWVEAGNFGQVNOF 180
Qy 456 KATCSGQTIKIDQKGSQIEPTPVIMTNNITVVRIGCEERPEHTQPIRDRMLNIHL 515
Db 181 KATCSGQTIKIDQKGSQIEPTPVIMTNNITVVRIGCEERPEHTQPIRDRMLNIHL 240
Qy 516 THLPGDFGLVDKNWPMICAWLVKQYQSTMASYCAKGVKPDWSENWAEKVPPTPINL 575
Db 241 VCKLPDGLVDKEEPMPLICAWLVKHYESTMANYTHHNGKVPEDENWAEKPIQEGVNS 300
Qy 576 LGSARSPFTTPKSTPLSQNYALTPLASDLELALPWSPTNPVAGTAEONTQNTGEAGSKA 635
Db 301 PG-CKOLETQASNPQSDHVLFLTPDVLVDALEPWSPTDPIETA-NQGSNQLG-VT 357
Qy 636 CQDQGLSPWSEIIBDLRACFGAEPLKQFSEPLN 670
Db 358 HKDQVQASPTWSEIIBDLRACFGAEPLKQFSEPLN 670

RESULT 11

S41439
gene NS-1 protein - Aleutian mink disease virus
C:Species: Aleutian mink disease virus
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S41439
R:Gottschalk, E.; Alexandersen, S.; Storgaard, T.; Bloom, M.E.; Aasted, B.

submitted to the EMBL Data Library, January 1994

A:Description: Sequence comparison of the non-structural genes of four different types of
A:Reference number: S41434
A:Accession: S41439
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-641 <GOT>
A:Cross-references: UNIPROT:Q65017; EMBL:Z29576
C:Genetics:
A:Introns: 586/1
C:Superfamily: parvovirus noncapsid protein

Query Match 15.1%; Score 550; DB 2; Length 641;
Best Local Similarity 25.8%; Pred. No. 6.3e-32;
Matches 178; Conservative 108; Mismatches 262; Indels 142; Gaps 25;

Qy 18 KEKSQEVFSFVFNENVOLNGKIDGWSYKKELEQDE-----LKSLOGRAETWD 68
Db 8 EORRLQDLYVLKKEIN--DGEVAVLFOQKTYTDDKONKPKATPPLTSSDLRLAFD 64
Qy 69 QSEDMEWETTVDWTK-----KQVFIQDSIAVKKCLEVLTNKNIFPDGVNWFVO 117
Db 65 SIEE-NLTASQLTNNNGINFCKLTLGLTKLLLDLKHVSHRWD--NNK-----VNLIWQ 115
Qy 118 HEMGKQDQWHCHVLIG-----GKDFSOAQGWRRQLN-----VYWSRWLVLTACNVQ- 164
Db 116 IEKGTQOQFHIHCCILGYPDNRNEDPKDVQKSLG-WFMKRLNKDLALIYNH-----HCDIQD 170
Qy 165 -LTPAERIKLRIADENNEWTLTKHKOTKDYKCVLFGNMIAYFYFKKKISTSPPR 223
Db 171 IKDPEDRAKNLVWVEDGPTKPYKFNQTKQDYKNPKVHLRDYTFIYLFNKKDKINTDSM- 229
Qy 224 DGYFLLSDSGWKTNFKLKEGRHLYSKLYTDDMRPETVETTV-----TTAQ 269
Db 230 -DGYFAAGGIVDN-LTNKERKTLRKMVLDQSDIDANDIDWEDGQDAPKVTDQDSA 287
Qy 270 ETKRG-----RIQTKKEVA-----IKTKELVHKRVTSPEWMMQ 307
Db 288 TTKTGTSLIWKSCATKVTSKKEVANPVQPSKKLYSAQNTLDALFNVGCFTEEDMIKQS 347
Qy 308 DSYIEMMAQPGENLLKNTLEICTITLARTKTAFLDILE--KAETSCLTNFSLPDTRICR 365
Db 348 DKYLSLEPFGQKINTLLHNNQVKTSTMTAFDCIIFKNEEEDDKPLATIKDM---- 403
Qy 366 IFAPFGWNVKVCIAICCVLRQGRKNTVLPFGPASTGKSIIAQAIAQAVGNVGVNAA 425
Db 404 -----GLNEQYLKVLCTILTKQGRKCIWYFGGCTGKTLASLI CKATVYGMVTS 458
Qy 426 NVNFPNDCTNKLIIWVEAGNFGQVNOFKAICSGQTIKIDQKGSQIEPTPVIMT 485
Db 459 NPNFPWTDGNNRIIWAEECGNLGNWVEDFKAITGGGVKVDTKNKKQPSIKGC-VVTS 517
Qy 486 NENITVVRIGCEERPEHTQPIRDRMLNIHL-----THLPGDFGLVDKNWPMICAWL 538
Db 518 NTNITKVTGVCEVTHAEPLKQRMKIRCMKTIINPKTKITFG-----MLKRWL 566
Qy 539 VKNQYQSTMASYCAKGVKVP-DWSENWAEKVPPTPINLLGSARSPFTTPKSTPLSQNYAL 597
Db 567 -----NTWDRQPIQLSHEMPELILET-----TGNSSATTATKNTGNSQ----- 605
Qy 598 TPLASDLELALPWSPTNPVAGTAEONTQ 627
Db 606 PTTAKSAESVNTNCDTPKRGASSVPPKQH 635

RESULT 12

S41434
gene NS-1 protein - Aleutian mink disease virus
C:Species: Aleutian mink disease virus
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S41434
R:Gottschalk, E.; Alexandersen, S.; Storgaard, T.; Bloom, M.E.; Aasted, B.
submitted to the EMBL Data Library, January 1994

A;Description: Sequence comparison of the non-structural genes of four different types of parvovirus
A;Reference number: S41434
A;Accession: S41434
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-641 <GOT>
A;Cross-references: UNIPROT:Q65020; EMBL:Z29577
C;Genetics:
A;Introns: 586/1
C;Superfamily: parvovirus noncapsid protein

Query Match 15.0%; Score 547.5; DB 2; Length 641;
Best Local Similarity 25.3%; Pred. No. 9.6e-32;
Matches 183; Conservative 109; Mismatches 260; Indels 171; Gaps 26;

Qy 18 KEKSNQEVFSFVKENNVQLNGDKIGWNSYKKEQLQEDE-----LKSQRGAETWTD 68
Db 8 EQRRLQELFE-KFKTE--VADGEGLAWLFQKQVTDKDNKPTKATPLRTTSSDLRLAFD 64
Qy 69 QSEDMEWETTVDKTKQVIFDLSLVKKCLFEVLNTKNI-----PGDVNMFVQHEWKGKD 123
Db 65 SIBE-TLKTSNQCLTNNDNFCKLTGLKTL--VLLDKHVRSHRDANKVNFVQVEKGKT 121
Qy 124 QGWHCHVLIG-----GKDFSAQAGKWRRLN-----VYMSRWLVACNVOLTPAERIK 172
Db 122 QQFHHCHCLGYFDKDEDSKDVQKSLG-WFKLNKOLAVIWSNHCDIQIGQSDGGRADN 180
Qy 173 LREIAEDNEWVLLTYKHKOTKDYTKCVLFGNMIAYFLTKKISTSPRDRG--GYFLS 230
Db 181 LKIWIEDGP-TKPKYKFNKQKDYKPNKPTNLRYLLYFNKDKIT-----QEGMDGYAS 235
Qy 231 SDSGWKNTPKGEERHLVSKLYTDDM-----RPETVETVTTAQTETKRG 274
Db 236 GNGGIIDN-LTNKERRALKRWYLDQSQGILDEDWEDSQSAPKVTDDQTSATSKTGS 294
Qy 275 -----RIQTKKEVA-----KTLKELVHKRVTSPEDEMMQSDSIEMM 314
Db 295 LVWKSATKVTSSKEVAIPVKQPSKQWTSQNTLDDLYMFCGCTPEDMLKMSDRYLEMS 354
Qy 315 AOPGGENLLKNTLEICTLTARTKTAFLDLE--KAETSKLTNFSLPDTRTCRIFAFHGW 372
Db 355 LEPNGAQKINTLLHMNQVTSNNITAFECIIKNEEBDEKPLDITKDM-----GL 405
Qy 373 NYVKVCHAICCVLNRQGGKNTVLFHGPASTGKSIIAQIAQAVNGVGVNAANVPFPN 432
Db 406 NEQHLKVLCTILTQSGKRGCVWFYFGPGTGKTLASLICKAVNYGMVTSNPNFPWT 465
Qy 433 DCTNKNLIWBEAGNFGQOVNQFKAICSQGTIRIDQKGSQKQIEPTPTVIMTNNITVV 492
Db 466 DCGNRNIIWAEEGNGNYYVEDFKAITGGDVKVDTKNKPQSIKGS-VIVTSNTNITKV 524
Qy 493 RIGCEERPEHTQPIRDRMLNIHL-----THLPGDFGLVDKNEWPMICAWLVKNGYQS 545
Db 525 TVGCVETNVHAEPLKQRMVKIRCKMPVNPKTTPG-----MLKTLI----- 566
Qy 546 TMASYCAKQKVPDWSNNAEPKVPTPINLLGSARSPTTPKSTPLSQNVALTPLASDLE 605
Db 567 -----STWDRIIP-----IKLSHEMP 581
Qy 606 DLALPEWSTNTVAGTAETONTGEAGSKACQDGLSPTWSEIEEDLRACFGAEPLKDF 665
Db 582 ELYLET-SGPNs--SATTATKST-----GNLQPTTAETAESVNTANCNDTPKRGAS 628
Qy 666 SEP 668
Db 629 SVP 631

RESULT 13
UYVPAP
noncapsid protein NS1 - Aleutian mink disease virus (strain ADV-G)
N;Alternate names: left-ORF protein
C;Species: Aleutian mink disease virus

C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C;Accession: A36760; A35529
R;Bloom, M.E.; Alexandersen, S.; Perryman, S.; Lechner, D.; Wolfbarger, J.B.
J. Virol. 62, 2903-2915, 1988
A;Title: Nucleotide sequence and genomic organization of Aleutian mink disease parvovirus
A;Reference number: A36760; MUID:88275062; PMID:2839709
A;Accession: A36760
A;Molecule type: DNA
A;Residues: 1-620 <BL2>
A;Cross-references: UNIPROT:P24030; EMBL:M20036
C;Superfamily: parvovirus noncapsid protein
C;Keywords: noncapsid protein

Query Match 14.9%; Score 543; DB 1; Length 620;
Best Local Similarity 26.6%; Pred. No. 1.9e-31;
Matches 157; Conservative 104; Mismatches 229; Indels 100; Gaps 18;

Qy 18 KEKSNQEVFSFVKENNVQLNGDKIGWNSYKKEQLQEDELKSLQGAETWTDQSE-DMEWE 76
Db 38 EQRRLQDLVYVOLKKEIN---DGEVAVLWFOQKTYTDKDNKPTKATPLRTTSSDLRLAFD 94
Qy 77 TTVDKMTKKQVFIQFISLVKKCLFEVLNTKNIF-----PGDVNMFVQHEWKGQDQGW 126
Db 95 SIBEENLTASNEHLTNNEINFCKLTGLKTLILLIDKHVSHRWDNSKNVLIWQIEKGTQOF 154
Qy 127 HCHVLIG-----GKDFSAQAGKWRRLN-----VYMSRWLVACNVQ--LTPAERIKL 173
Db 155 HIHCLGLGYFDKDNBDKDVQKSLG-WFMKRLNKLDAVIYSN---HHCDIQIDKDPEDRAKN 210
Qy 174 RETAEDNEWVLLTYKHKOTKDYTKCVLFGNMIAYFLTKKISTSPRDRGCVFLSSDS 233
Db 211 LKVMIEDGPTKPKYKFNKQKDYKPNKPTNLRYLLYFNKDKINTDSM--DGYFAAGNG 269
Qy 234 GWKTNFLKEGERHLVSKLYTDDMRPTEVTV-----TTAQETKRG----- 274
Db 269 GIVDN-LTNKERTKRLKWLDEQSSDIMDANIDWEDQDAPKVTDDQTSATTKGTSLIW 327
Qy 275 -----RIQTKKEVA-----IKTTLKELVHKRVTSPEDEMMQSDSIEMMAQP 317
Db 328 KSCATKVTSSKEVANVPQPSKKLYSAQSTLDALFNVCFTPEDMIKQSDKYLESLSEP 387
Qy 318 GGENLLKNTLEICTLTARTKTAFLDLE--KAETSKLTNFSLPDTRTCRIFAFHGWYV 375
Db 388 NGPQKINTLLHMNQVTSNTITAFDCIIFNEEBDDKPLLATIKDM-----GLNEQ 438
Qy 376 KVCHAICCVLNRQGGKNTVLFHGPASTGKSIIAQIAQAVNGVGVNAANVPFPNDCT 435
Db 439 YLKKVLCITLTQSGKRGCVWFYFGPGTGKTLASLICKATVNYGMVTSNPNFPWTDG 498
Qy 436 NKNLIIWBEAGNFGQOVNQFKAICSQGTIRIDQKGSQKQIEPTPTVIMTNNITVVIRIG 495
Db 499 NRNIIWAEEGNGNFWVEDFKAITGGDVKVDTKNKPQSIKGC-VIVTSNTNITKVTVG 557
Qy 496 CEERPEHTQPIRDRMLNIHL-----THLPGDFGLVDKNEWPMICAWL 538
Db 558 CVETNAHAEPFLKQRMKIRCKMTINPKTITPG-----MLKRWL 596

RESULT 14
S41861
gene NS-1 protein - Aleutian mink disease virus
C;Species: Aleutian mink disease virus
C;Date: 15-Sep-1994 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
C;Accession: S41861; S41436
R;Gottschalk, E.; Alexandersen, S.; Storgaard, T.; Bloom, M.E.; Aasted, B.
submitted to the EMBL Data Library, January 1994
A;Description: Sequence comparison of the non-structural genes of four different types of
A;Reference number: S41434
A;Accession: S41861
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-641 <GOT>
A;Cross-references: UNIPROT:Q65023; EMBL:X77085; NID:g452597; PIDN:CAA54355.1; PID:g45259

C;Genetics:

A; Introns: 586/1

C; Superfamily: parvovirus noncapsid protein

Query Match 14.7%; Score 533.5; DB 2; Length 641;
Best Local Similarity 25.7%; Pred. No. 1e-30;
Matches 178; Conservative 106; Mismatches 250; Indels 159; Gaps 27;

QY 18 KEKSNQVFSVFVNENVLQNGKDGWNSYKKEIQEDE-----LKSILQGAETWD 68
DB 8 EQRKLQDIYQLQKE---VADGEGLAWLFOQKYTTDKNRPATPPLRTSSDLRLAFD 64
QY 69 QSEDM-----KKQVFIEDSLVKKCLFEVLNFKNIFPDPVNVFVQ 117
DB 65 SIBETLIASNEW-LTKDEINFCKLTGLKTLVLVDKHVKSHRWA-----DKINFIMQ 115
QY 118 HENWKGQGWCHVLIG-----GKFSQAQGWRRQLNYYWSRMLVTCNVQLTPAER 170
DB 116 IEKGTQHFHHCGLGYFDKNEPKQVQKSLG-WLIKLN-----RDLAATFSNHHCDQD 170
QY 171 IKLREIAEDN--EWV-----TLTYKHQTKDKYTKCVLFGNMIAYVFLTKKISTSPPR 223
DB 171 IKOPEAKANLKWIEDGPTKPKYNNKQTKDYKRPVTLSDVYIYLFNKDIIH-----K 226
QY 224 DG--GYPLSSDSGKTNFLKEGERHLVSKLYTDDMRPETVE-----TTVTT 267
DB 227 EGMGYYAAGNGGLIDN-LTNKERRKALRKNYLDQSSDIMDADIDWEDGQAPKVDQTD 285
QY 268 AQETKRG-----RIOTKKEVA-----IKTTLKELVHKRVTSPEWMM 305
DB 286 SATSKTSTLIWKSATKVTSTKEVANPVQOPSKKLYSAQNTLDALFNVCFTPEDMIK 345
QY 306 QPDSYIEMMAQPGENLLKNTLEICTLTARTKTAFDLILE--KAETSKLTNFSLPDTRT 363
DB 346 QSDKYLELSELPNGPKINTLLHNVQKSTMTMTAFDCIIFNEBEDDKPLATIIDM-- 403
QY 364 CRIFAFHGMNVKCHAI CCVLNRQGGKRNVLPHGPASTGKSIIAQIAVGNVCYN 423
DB 404 -----GLNEQYLKVKVLTILTQGGKRCIWFYGGTGKTLASLICKATVNYGMVT 456
QY 424 AANVPFPNDCTNKLIVWEAGNFQGVNQFKAICSGQTIRIDQKGSQKIEPTPVIM 483
DB 457 TSNPFPFTDCGNRIIWAEECNLGNWVEDFKAITGGGDKVYDTKNKQPSIKGC-VIV 515
QY 484 TTINENTIVVRIGCEERPEHTQPIRDRMLNIHL-----THLPDGFGLVDKNWPMICA 536
DB 516 TSNNTIKVTGCVETNAHAEPKQRMKIRCMKTINPKTITPG-----MLKR 564
QY 537 WLKNGVQSTMASYCAKWKVP-----DWSNWAEPKVPTPINLLGSARSPP-TTPKSTP 590
DB 565 WL-----NTWDRQPTQLSHEMPELYLETTGPN-----SSATSATKTTGNSQP 606
QY 591 LSQNYALTPLASDLELPEWSTPNTPVAGTA 623
DB 607 TTAETAESVSTADCD-----TPKRGAS 628

RESULT 15

S52209

noncapsid protein NS1 - muscovy duck parvovirus

N; Alternate names: nonstructural protein NS1; NS1 protein

C; Species: muscovy duck parvovirus

C; Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004

C; Accession: S52209

R; Zadori, Z.; Erdei, J.; Nagy, J.; Kisari, J.

submitted to the EMBL Data Library, September 1993

A; Reference number: S52209

A; Accession: S52209

A; Molecule type: DNA

A; Residues: 1-626 <ZAD>

A; Cross-references: UNIPROT:Q83288; EMBL:X75093; NID:g609091; PIDN:CAA52983.1; PID:g60909

A; Experimental source: strain FM

C; Genetics:

A; Gene: NS1

C; Superfamily: parvovirus noncapsid protein

C; Keywords: noncapsid protein

Query Match 12.4%; Score 452.5; DB 2; Length 626;
Best Local Similarity 25.2%; Pred. No. 7.5e-25;
Matches 164; Conservative 81; Mismatches 265; Indels 141; Gaps 22;

QY 75 WETT-----VDEMTKQVFI FDSILVKKCLFEVLNKNIF-----PGDVNFVQHEWGK 122
DB 45 WBEPTGIWNNHEVNLPMVTLADKI-----KNIFQIRWNNQFNQDRTDFFQLEEGS 93
QY 123 DQGWCHVLIGGKDFSAQGWRRQLNYYWSRMLVTCNVQLTPAERIK--LREIAD 179
DB 94 EY-IHLHACVPGECRSFVLGRY-----MSQIKDSILRDVYEG 129
QY 180 N-----EWTLTYKHQTKDKYTKCVLFGNMIAYVFLTKKISTSPPRDGGYFSSDSG 234
DB 130 KQVKIPDMFSTIKTKRGQGNKTVTAA-----YILHYLIPKQ-----PELQWAFNTMPLF 179
QY 235 WKTNFKLKEGERHLVSKLYTDDMRPETVETVTTAQETKEGRITQTKKEVAIKTTLKELVHK 294
DB 180 TTAALCLQKRGELLDAFQSENNVAVQEDQAATAPL-----ISNRAAKNYSNLVDWLIEM 235
QY 295 RVTSPEDWMMOPDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLT 354
DB 236 GITSEKQWLTKENKESYRSFQATSSNNRQVKAALENARAEMLLTKTATDYLIGKDPVLDIT 295
QY 355 NPSLPDTRTCRIFAFHGMNVKCHAI CCVLNRQGGKRNVLPHGPASTGKSIIAQIAO 414
DB 296 K-----NRIYQILKLNYPQVGVLCGWVKRBEFNKRNAILWYGPATTTGKTNIAEATIAH 350
QY 415 AVGNVCYNAANVPENDCTNKLIVWEAGNFQGVNQFKAICSGQTIRIDQKGSK 474
DB 351 AVPFYGCNVNNTNENPFNDVCDKMLIWEEGKNTKRVESAKAILGGSNAVRVDQCKGVS 410
QY 475 QIEPTPVIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGFGLVDKN----- 530
DB 411 CIEPTPVITSTNDMCMIVDGNSTTMEHRIPLEERMFQIVLSHKLGNFGKISKEVKEF 470
QY 531 --W-----PMTCAWLKNGYQSTWASYCAKWKGVDPDWSN-----WAEPK----- 568
DB 471 FKWANDNLVPVSEFQKVPNTNEQTKLTE-----PVPERANEFSEPPPKI WAPTRREEBEEI 524
QY 569 -----VPTPINLLGSARSPTTPKSTPLSQNYALTPLASDLELDALEPMSTENT 617
DB 525 LRASPELFASVAPLP-----SSPDTSEKPKKTRGETQVRCAMHSL-DNSMNVFECLC 576
QY 618 PVAGTAETQNTGEAGSKAC-QDQGLSPTWSIEEDLDRACFGAEPKPKDFSE 667
DB 577 ERANPPEFQSLGE---NFCNQH-----WYD-----CAFCELNKDDWNE 612

Search completed: January 22, 2005, 03:41:57

Job time : 26.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:19:28 ; Search time 106.333 Seconds
(without alignments)
3636.223 Million cell updates/sec

Title: US-10-069-056-6
Perfect score: 3638
Sequence: 1 MAGNAYSDEVLGATNWLKKE.....RACFGAEPLKDFSEPLNLD 672

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_02.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3625	99.6	672	1	VNCS_MUMIV	P03134 murine minu
2	3625	99.6	721	2	Q84365	Q84365 murine minu
3	3556	97.7	721	1	VNCS_MUMIM	P07300 murine minu
4	3556	97.7	721	2	Q84363	Q84363 murine minu
5	3554	97.7	672	2	Q83429	Q83429 mouse parvo
6	3347	92.0	672	2	Q8JV14	Q8JV14 rat minute
7	3338	91.8	672	2	Q8JV28	Q8JV28 kilham rat
8	3337	91.7	672	2	P88899	P88899 kilham rat
9	3329	91.5	672	2	Q8JV18	Q8JV18 rat minute
10	3328	91.5	672	1	VNCS_PAVHH	P03133 hamster par
11	3325	91.4	672	2	Q8JV16	Q8JV16 rat minute
12	3324	91.4	668	1	VNCS_PAVL3	P36311 parvovirus
13	3305	90.8	665	2	O71159	O71159 kilham rat
14	2937.5	80.7	671	2	O71157	O71157 rat parvovi
15	2657	73.0	668	2	P89513	P89513 feline panl
16	2656	73.0	668	2	P89516	P89516 feline panl
17	2655	73.0	668	2	P89512	P89512 feline panl
18	2655	73.0	668	2	P89515	P89515 feline panl
19	2655	73.0	668	2	P90449	P90449 feline panl
20	2655	73.0	668	2	P90484	P90484 feline panl
21	2654	73.0	668	2	P89514	P89514 feline panl
22	2654	73.0	668	2	P90472	P90472 feline panl
23	2649	72.8	668	1	VNCS_MEVA	P27438 mink enteri
24	2646	72.7	668	1	VNCS_FPV19	P24842 feline panl
25	2644	72.7	668	1	VNCS_PAVCN	P12929 canine parv
26	2639	72.5	668	2	Q84393	Q84393 canine parv
27	2616	71.9	668	2	Q70M74	Q70M74 canine parv
28	2615	71.9	668	2	CAE47433	CAE47433 canine pa
29	2487	68.4	660	1	VNCS_PAVPN	P18547 porcine par
30	2478	68.1	662	1	VNCS_PAVPK	P52502 porcine par
31	2474	68.0	662	2	Q6RED5	Q6RED5 porcine par

RESULT 1

ID	VNCS_MUMIV	STANDARD;	PRT;	672 AA.
AC	P03134;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVP1).			
GN	Name=NS1;			
OS	Murine minute virus (Murine parvovirus).			
OC	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.			
OX	NCBI_TaxID=10794;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=83143341; PubMed=6298737;			
RA	Astell C.R., Thomson M., Merchlinsky M., Ward D.C.;			
RT	"The complete DNA sequence of minute virus of mice, an autonomous parvovirus.";			
RL	Nucleic Acids Res. 11:999-1018(1983).			
RN	[2]			
RP	HOMOLOGIZATION, AND INTERACTION WITH SYNCRIP.			
RP	MEDLINE=99102562; PubMed=9847309;			
RA	Harris C.E., Boden R.A., Astell C.R.;			
RT	"A novel heterogeneous nuclear ribonucleoprotein-like protein interacts with NS1 of the minute virus of mice.";			
RL	J. Virol. 73:72-80(1999).			
CC	-I- FUNCTION: Seems necessary for viral DNA replication.			
CC	-I- SUBUNIT: Forms homooligomers. Interacts with human SYNCRIP.			
CC	-I- DOMAIN: The N-terminus (residues 1-275) possess a negative effect on transactivation.			
CC	-I- DOMAIN: The C-terminus (residues 543-672) possess an activation domain.			
CC	-I- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; J02275; AAA67109.1; -			
DR	EMBL; V01115; CAA24309.1; ALT_INIT.			
DR	PIR; A03696; UYPVIM.			
DR	TRANSFAC; T02375; -			
DR	InterPro; IPR001257; Parvo NS1.			
DR	PFam; PF01057; Parvo NS1; I.			
KW	ATP-binding; DNA replication; Noncapsid protein;			
KW	Nonstructural protein.			
FT	DOMAIN 1 276 Interacts with SYNCRIP.			
FT	NP_BIND 399 406 ATP (Potential)			
SQ	SEQUENCE 672 AA; 76248 MW; 50298F27662E3C1D CRC64;			

O6TPD8 porcine par
Aa90279 porcine p
Aa91039 porcine p
Q6PS60 porcine par
Aa93262 porcine p
Q993m6 autonomous
P06431 feline panl
Q65017 aleutian mi
Q65020 aleutian mi
P24030 aleutian mi
Q96607 aleutian mi
Q65023 aleutian mi
Q6J180 bovine aden
Aa26464 bovine ad

```
Query Match 99.6%; Score 3625; DB 1; Length 672;
Best Local Similarity 99.7%; Pred. No. 1.e-260;
Matches 670; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVIGATNWLKESNQVFSFVKNNVQNGKDIGWNSYKKELOEDELKSLQ 60
DB 1 MAGNAYSDEVIGATNWLKESNQVFSFVKNNVQNGKDIGWNSYKKELOEDELKSLQ 60
QY 61 RGAETTTWQSDMEWETTVDEMTKKQVIFDLSVKKCLFEVLNTKNIFFPGDVNWFVQHEW 120
DB 61 RGAETTTWQSDMEWETTVDEMTKKQVIFDLSVKKCLFEVLNTKNIFFPGDVNWFVQHEW 120
QY 121 GKDGQWHCHVLIGKDFSOAQGWRRQLNVYWSRWLVTAACNVQLTPAERIKLREIAEDN 180
DB 121 GKDGQWHCHVLIGKDFSOAQGWRRQLNVYWSRWLVTAACNVQLTPAERIKLREIAEDN 180
QY 181 EWTLLTYKHQTKDQYTKCVLFGNMIAYFLTKKISTSPRRGGYFLSSDGSWKTNFL 240
DB 181 EWTLLTYKHQTKDQYTKCVLFGNMIAYFLTKKISTSPRRGGYFLSSDGSWKTNFL 240
QY 241 KEGERHLVSKLYTDDMRPETVETTTAQTETKRGRIQTKKEVAIKTTLKELVHKRVTSPE 300
DB 241 KEGERHLVSKLYTDDMRPETVETTTAQTETKRGRIQTKKEVAIKTTLKELVHKRVTSPE 300
QY 301 DWMMPQDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSKLTNFSLPD 360
DB 301 DWMMPQDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSKLTNFSLPD 360
QY 361 TRTCRIFAFHGNVYKVAICCVLNROGGKNTVLFHGPASTGKSIITAAIAQAQVGNVG 420
DB 361 TRTCRIFAFHGNVYKVAICCVLNROGGKNTVLFHGPASTGKSIITAAIAQAQVGNVG 420
QY 421 CYNAAVNFPPNDCTNKNLIWVEEAGNFQGVNQVQFKAICSGQTIRIDQKGSQKQIEPTP 480
DB 421 CYNAAVNFPPNDCTNKNLIWVEEAGNFQGVNQVQFKAICSGQTIRIDQKGSQKQIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
DB 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
QY 541 NGYSTWASVYCAKWKGVDPDSENWAEKVPPTPINLGSARSPTTTPKSTPLSONYALTPL 600
DB 541 NGYSTWASVYCAKWKGVDPDSENWAEKVPPTPINLGSARSPTTTPKSTPLSONYALTPL 600
QY 601 ASDLEDLALPEWSTPTNPVAGTAETQNTGEAGSKACQDQGLSPTWSEIEEDLRACFGAEP 660
DB 601 ASDLEDLALPEWSTPTNPVAGTAETQNTGEAGSKACQDQGLSPTWSEIEEDLRACFGAEP 660
QY 661 LKXDFSEPLNLD 672
DB 661 LKXDFSEPLNLD 672

RESULT 2
Q84365 PRELIMINARY; PRT; 721 AA.
AC Q84365;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nonstructural protein.
GN Name=NSI;
OS Murine minute virus (Murine parvovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MVM;
RX MEDLINE=83143341; PubMed=6298737;
RA Astell C.R., Thomson M., Merchlinsky M., Ward D.C.;
RT "The complete DNA sequence of minute virus of mice, an autonomous
parvovirus.";
RL Nucleic Acids Res. 11:999-1018(1983).
```

```
QY 661 LKXDFSEPLNLD 672
Db 710 LKXDFSEPLNLD 721

RESULT 3
VNC5 MUMIM STANDARD; PRT; 672 AA.
AC P07300; P10837;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVP1).
GN Name=NS1;
OS Murine minute virus (strain MVM1) (Murine parvovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OC NCBI_TaxID=10795;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86115415; PubMed=3502703;
RA Astell C.R., Gardiner E.M., Tattersall P.;
RT "DNA sequence of the lymphotropic variant of minute virus of mice,
RT MVM(1), and comparison with the DNA sequence of the fibrotropic
RT prototype strain.";
RL J. Virol. 57:656-669(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85242059; PubMed=3855242;
RA Sahli R., McMaster G.K., Hirt B.;
RT "DNA sequence comparison between two tissue-specific variants of the
RT autonomous parvovirus, minute virus of mice.";
RL Nucleic Acids Res. 13:3617-3633(1985).
CC -!- FUNCTION: Seems necessary for viral DNA replication.
CC -!- SUBUNIT: Forms homooligomers. Interacts with human SYNCRIP (By
CC similarity).
CC -!- DOMAIN: The N-terminus (residues 1-275) possess a negative effect
CC on transactivation (By similarity).
CC -!- DOMAIN: The C-terminus (residues 543-672) possess an activation
CC domain (By similarity).
CC -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02481; -; NOT ANNOTATED_CDS.
DR EMBL; M12032; AAA69567.1; -.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; 1.
KW ATP-binding; DNA replication; Noncapsid protein;
KW Nonstructural protein.
FT DOMAIN 1 276 Interacts with SYNCRIP (By similarity).
FT NP BIND 399 406 ATP (Potential).
FT FT CONFLICT 597 597 I -> L (in Ref. 2).
SQ SEQUENCE 672 AA; 76140 MW; 25F025FE328B4DF0 CRC64;

Query Match 97.7%; Score 3556; DB 1; Length 672;
Best Local Similarity 97.5%; Pred. No. 1.4e-255;
Matches 655; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVILGATNMLKESKNOEVSFVKPNVOLNGDKIGWNSYKKELODELKSLQ 60
Db 1 MAGNAYSDEVILGATNMLKESKNOEVSFVKPNVOLNGDKIGWNSYKKELODELKSLQ 60
QY 61 RGAETTWQSQDEMEWETTVDMTKKQVFIPIFDSLVKKCLFEVLNTKNIIFPGDVNWFVQHEW 120
Db 61 RGAETTWQSQDEMEWETTVDMTKKQVFIPIFDSLVKKCLFEVLNTKNIIFPGDVNWFVQHEW 120
QY 121 GKDQGHCHVLIGGKDFSOAQGWRRQLNVYWSRWLVTAACNVQLTPAERIKLUREIAEDN 180
```

```
Qy 1 MAGNAYSDEVLGATNWLKEKSNQEVFSFVKENVLQNGKDIGNWSYKKELOEDELKSLQ 60
Db 50 MAGNAYSDEVLTGTTNWLKEKSNQEVFSFVKETEDVLQNGKDIGNWYKKELOEDELKSLQ 109
Qy 61 RGAETTWDOSEDMEWETTVDKTKQVFIYDLSLVKKCLFEVLNTKNIIFPGDVNWFVQHEW 120
Db 110 RGAETTWDOSEDMEWETTVDKTKQVFIYDLSLVKKCLFEVLNTKNIAPADVTWVQHEW 169
Qy 121 GKDOGWCHVLIGGKDFSAQGWRRQLNVMYWSRLVLTACNVOLTPAERIKLREIAEDN 180
Db 170 GKDOGWCHVLIGGKDFSAQGWRRQLNVMYWSRLVLTACNVOLTPAERIKLREIAEDS 229
Qy 181 EWTLLTYKHQTKKDYTKCVLFGNMIAIYFLLTKKISTGPPRDGGYFLSSDGSWKTNFL 240
Db 230 EWTLLTYKHQTKKDYTKCVLFGNMIAIYFLLTKKISTGPPRDGGYFLSSDGSWKTNFL 289
Qy 241 KEGERHLVSKLYTDDMRPETVETVTTAQTETKRGRIQTKKEVAIKTTLKELVHKRVTSPE 300
Db 290 KEGERHLVSKLYTDDMRPETVETVTTAQTETKRGRIQTKKEVAIKTTLKELVHKRVTSPE 349
Qy 301 DMMWQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAEYSKLTNFSPLD 360
Db 350 DMMWQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAEYSKLTNFSPLD 409
Qy 361 TRTCRIFAFHGWNYVKVCHAI CCVLNRQGGKNTVLPHGPASTGKSI IAQAIAQAVGNVG 420
Db 410 TRTCRIFAFHGWNYVKVCHAI CCVLNRQGGKNTVLPHGPASTGKSI IAQAIAQAVGNVG 469
Qy 421 CYNAAVNFPPNDCTNKNLIWVEAGNFGQVNOFKAI CSQGTIRIDOKGSKQIEPTP 480
Db 470 CYNAAVNFPPNDCTNKNLIWVEAGNFGQVNOFKAI CSQGTIRIDOKGSKQIEPTP 529
Qy 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMI CAWLK 540
Db 530 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMI CAWLK 589
Qy 541 NGYSTWASACAKGKVPDWSNWAEPKVPPTPINLGSARSPTTPKSTPLSQNYALTPL 600
Db 590 NGYSTWASACAKGKVPDWSNWAEPKVPPTPINLGSARSPTTPKSTPLSQNYALTPL 649
Qy 601 ASDLEDLALPWSPTNPVAGTAETONTGEAGSKACODGOLSPTWSEIEEDLRACFGAEP 660
Db 650 ASDLEDLALPWSPTNPVAGTAETONTGEAGSKACODGOLSPTWSEIEEDLRACFGAEP 709
Qy 661 LKPDFSEPLNLD 672
Db 710 LKPDFSEPLNLD 721

RESULT 5
Q83429 ID Q83429 PRELIMINARY; PRT; 672 AA.
AC Q83429
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nonstructural protein 1.
GN Names=NS1;
OS Mouse parvovirus 1.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=35340;
RN [1]
RP MEDLINE=94365951; PubMed=8083985;
RX Ball-Goodrich L.J., Johnson B.;
RA "Molecular characterization of a newly recognized mouse parvovirus.";
RL J. Virol. 68:6476-6486(1994).
DR EMBL; U12469; AAA61405.1; --
DR GO; GO:0019012; C:virion; IEA.
DR GO; GO:0019079; P:Viral genome replication; IEA.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; I.
KW Nonstructural protein.
```

```
SQ SEQUENCE 672 AA; 76112 MW; 31CG365276727363 CRC64;
Query Match 97.7%; Score 3554; DB 2; Length 672;
Best Local Similarity 97.3%; Pred. No. 1.9e-255;
Matches 654; Conservative 11; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MAGNAYSDEVLGATNWLKEKSNQEVFSFVKENVLQNGKDIGNWSYKKELOEDELKSLQ 60
Db 1 MAGNAYSDEVLTGTTNWLKEKSNQEVFSFVKETEDVLQNGKDIGNWYKKELOEDELKSLQ 60
Qy 61 RGAETTWDOSEDMEWETTVDKTKQVFIYDLSLVKKCLFEVLNTKNIIFPGDVNWFVQHEW 120
Db 61 RGAETTWDOSEDMEWETTVDKTKQVFIYDLSLVKKCLFEVLNTKNIAPADVTWVQHEW 120
Qy 121 GKDOGWCHVLIGGKDFSAQGWRRQLNVMYWSRLVLTACNVOLTPAERIKLREIAEDN 180
Db 121 GKDOGWCHVLIGGKDFSAQGWRRQLNVMYWSRLVLTACNVOLTPAERIKLREIAEDS 180
Qy 181 EWTLLTYKHQTKKDYTKCVLFGNMIAIYFLLTKKISTGPPRDGGYFLSSDGSWKTNFL 240
Db 181 EWTLLTYKHQTKKDYTKCVLFGNMIAIYFLLTKKISTGPPRDGGYFLSSDGSWKTNFL 240
Qy 241 KEGERHLVSKLYTDDMRPETVETVTTAQTETKRGRIQTKKEVAIKTTLKELVHKRVTSPE 300
Db 241 KEGERHLVSKLYTDDMRPETVETVTTAQTETKRGRIQTKKEVAIKTTLKELVHKRVTSPE 300
Qy 301 DMMWQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAEYSKLTNFSPLD 360
Db 301 DMMWQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAEYSKLTNFSPLD 360
Qy 361 TRTCRIFAFHGWNYVKVCHAI CCVLNRQGGKNTVLPHGPASTGKSI IAQAIAQAVGNVG 420
Db 361 TRTCRIFAFHGWNYVKVCHAI CCVLNRQGGKNTVLPHGPASTGKSI IAQAIAQAVGNVG 420
Qy 421 CYNAAVNFPPNDCTNKNLIWVEAGNFGQVNOFKAI CSQGTIRIDOKGSKQIEPTP 480
Db 421 CYNAAVNFPPNDCTNKNLIWVEAGNFGQVNOFKAI CSQGTIRIDOKGSKQIEPTP 480
Qy 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMI CAWLK 540
Db 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMI CAWLK 540
Qy 541 NGYSTWASACAKGKVPDWSNWAEPKVPPTPINLGSARSPTTPKSTPLSQNYALTPL 600
Db 541 NGYSTWASACAKGKVPDWSNWAEPKVPPTPINLGSARSPTTPKSTPLSQNYALTPL 600
Qy 601 ASDLEDLALPWSPTNPVAGTAETONTGEAGSKACODGOLSPTWSEIEEDLRACFGAEP 660
Db 601 ASDLEDLALPWSPTNPVAGTAETONTGEAGSKACODGOLSPTWSEIEEDLRACFGAEP 660
Qy 661 LKPDFSEPLNLD 672
Db 661 LKPDFSEPLNLD 672

RESULT 6
Q8JV14 ID Q8JV14 PRELIMINARY; PRT; 672 AA.
AC Q8JV14;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nonstructural protein 1.
GN Name=NS1;
OS Rat minute virus 1c.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=172387;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2121070; PubMed=12124471;
RA Wan C.H., Soderlund-Venermo M., Pintel D.J., Riley L.K.;
RT "Molecular characterization of three newly recognized rat parvoviruses.";
KW
```



```
RL J. Gen. Virol. 83:2075-2083 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Wan C.-H., Pintel D.J., Soderlund-Venermo M., Riley L.K.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF332884; AAM93279.1; -.
DR GO: GO:0019012; C:virion; IEA.
DR GO: GO:0019079; P: viral genome replication; IEA.
DR InterPro: IPR001257; Parvo_NSL.
DR Pfam: PF01057; Parvo_NSL; 1.
KW Nonstructural protein.
SQ SEQUENCE 672 AA; 75988 MW; 52DF654934CF3FD CRC64;

Query Match 92.0%; Score 3347; DB 2; Length 672;
Best Local Similarity 91.7%; Pred. No. 4.9e-240;
Matches 616; Conservative 23; Mismatches 33; Indels 0; Gaps 0;

QY 1 MAGNAYDEVLGATNWLKESNQEVFVFNENVQNGKDIGNSYKKELODELKSLQ 60
DB 1 MAGNAYDEVLGATNWLKSSQEVFVFNENVQNGKDIGNSYKKELODELKSLQ 60
QY 61 RGAETTWQSDSEMEWETTVDEMTHKQVIFDLSLVKKCLFEVLNTKNIFFPGDVNMFVQHEW 120
DB 61 RGAETTWQSDSEMEWESAVIDMTKQVIFDLSLVKKCLFEVLSTKNIAPSDVTWVQHEW 120
QY 121 GKQGHCHVLIGKDFSOAQGWRRQLNYYWSRWLVTAQNVQLTPAERI KUREIAEDN 180
DB 121 GKQGHCHVLIGKDFSOAQGWRRQLNYYWSRWLVTAQNVQLTPAERI KUREIAEDS 180
QY 181 EWTLLTYKHQTKDYKCVLFGNMIAYVFLTKKISTSPRDGGVFLSSDGSWKTNFL 240
DB 181 EWTLLTYKHQTKDYKCVLFGNMIAYVFLSKKICTSPRDGGVFLSSDGSWKTNFL 240
QY 241 KEGERHLVSKLYTDDMRPETVETVTTAQTKEGRIOQTKKEVAIKTTLKELVHKRVTSPE 300
DB 241 KEGERHLVSKLYTDEMKPETVETVTTAQAEGRIQTRKEVSIKTTLLKELVHKRVTSPE 300
QY 301 DNMWQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360
DB 301 DNMWQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLANFSMAS 360
QY 361 TRTCRIFAEHGMNYYKCHAI CCVLRNQGKRNVLPHGPASTGKSI IAQIAQGVNNG 420
DB 361 TRTCRIFAEHGMNYYKCHAI CCVLRNQGKRNVLPHGPASTGKSI IAQIAQGVNNG 420
QY 421 CYNAANVNFNDCTNKNLIWVEEAGNFQGVNQPKAI CSQTIRIDQKGSQKQIEPTP 480
DB 421 CYNAANVNFNDCTNKNLIWVEEAGNFQGVNQPKAI CSQTIRIDQKGSQKQIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
DB 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
QY 541 NGYQSTWASYCAKWKVPDSENAEPKVPPTPINLGSARSPTTTPKSTPLSONYALTPL 600
DB 541 NGYQSTWASYCAKWKVPDSENAEPKVPPTPINLGSARSPTTTPKSTPLSONYALTPL 600
QY 601 ASDLEDLALPEWSTPNTPVAGTAETQNTGEAGSKACODGQSLSPTWSEIEDLACFCAEP 660
DB 601 ASDLEDLALPEWSTPNTPVAGTAETQNTGEAGSKACODGQSLSPTWSEIEDLACFCAEP 660
QY 661 LKXDFSEPLNLD 672
DB 661 LKXDFSEPLNLD 672

RESULT 7
Q8JV28 PRELIMINARY; PRT; 672 AA.
ID AC
IC Q8JV28:
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
```

```
Db      661 LKXDFSEPLNLD 672
       1:||||:|
       661 LESDFNEELTLD 672

RESULT 8
P88899 PRELIMINARY; PRT; 672 AA.
AC P88899;
DT 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Non-capsid protein.
GN Name=NS1;
OS Kilham rat virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=12441;
RN [1]
RP SEQUENCE FROM N.A.
RA Brown D.W., Like A.A.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U79033; AAB38326.1; -.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR InterPro; IPR001257; Parvo_NS1.
DR Pfam; PF01057; Parvo_NS1; 1.
SQ SEQUENCE 672 AA; 76056 MW; 0299B36871A0A10A CRC64;

Query Match 91.7%; Score 3337; DB 2; Length 672;
Best Local Similarity 91.4%; Pred. No. 2.7e-239;
Matches 614; Conservative 23; Mismatches 35; Indels 0; Gaps 0;

Qy 1 MAGNAYSDVLGATNWLKEKSNQVFSFVKFNENVLNGKDIGNWSYKKELODELKSLQ 60
Db 1 MAGNAYSDVLGATNWLKSSQSEVFSFVKFNENVLNGKDIGNWSYKKELODELKSLQ 60
Qy 61 RGAETTWDSQSDMEWETTVDMTKKQVIFDLSVKKCLFEVLNTKNIIPGDVNVFQHEW 120
Db 61 RGAETTWDSQSDMEWESAVDDMTKKQVIFDLSVKKCLFEVLNTKNIIPGDVNVFQHEW 120
Qy 121 GKDOGWCHVLIGGKDFSQAGKWRRLNVLNYSRWLVTAACNVLTTPAERIKLREIAEDN 180
Db 121 GKDOGWCHVLIGGKDFSQAGKWRRLNVLNYSRWLVTAACNVLTTPAERIKLREIAEDN 180
Qy 181 EWTLLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPRGGYFLSSDGSWKTNFL 240
Db 181 EWTLLTYKHQTKDYTKCVLFGNMIAYFLSKKICTSPRGGYFLSSDGSWKTNFL 240
Qy 241 KEGERHLVSKLYTDDMPETVTTTAQETKGRIOTKKEVAIKTTLKELVHKRVTSPE 300
Db 241 KEGERHLVSKLYTDEMPEVTTTVAQAKGRIOTKRKEVSIKTLKELVHKRVTSPE 300
Qy 301 DWMNQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSKLTNFSLPD 360
Db 301 DWMNQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSKLTNFSLPD 360
Qy 361 TRTCRIFAFHGWNVYKCHAI CCVLNRQGGKRNVLPHGPASTGKSIIAQIAQAVGNVG 420
Db 361 TRTCRIFAEHGWNVYKCHAI CCVLNRQGGKRNVLPHGPASTGKSIIAQIAQAVGNVG 420
Qy 421 CYNAAANVFPNDCTNKNLIWVEEAGNFQGVNQFKAICSGQTIIRDQKKGSKQIEPTP 480
Db 421 CYNAAANVFPNDCTNKNLIWVEEAGNFQGVNQFKAICSGQTIIRDQKKGSKQIEPTP 480

Query Match 91.5%; Score 3329; DB 2; Length 672;
Best Local Similarity 91.2%; Pred. No. 1.1e-238;
Matches 613; Conservative 24; Mismatches 35; Indels 0; Gaps 0;

Qy 1 MAGNAYSDVLGATNWLKEKSNQVFSFVKFNENVLNGKDIGNWSYKKELODELKSLQ 60
Db 1 MAGNAYSDVLGATNWLKSSQSEVFSFVKFNENVLNGKDIGNWSYKKELODELKSLQ 60
Qy 61 RGAETTWDSQSDMEWETTVDMTKKQVIFDLSVKKCLFEVLNTKNIIPGDVNVFQHEW 120
Db 61 RGAETTWDSQSDMEWESAVDDMTKKQVIFDLSVKKCLFEVLNTKNIIPGDVNVFQHEW 120
Qy 121 GKDOGWCHVLIGGKDFSQAGKWRRLNVLNYSRWLVTAACNVLTTPAERIKLREIAEDN 180
Db 121 GKDOGWCHVLIGGKDFSQAGKWRRLNVLNYSRWLVTAACNVLTTPAERIKLREIAEDN 180
Qy 181 EWTLLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPRGGYFLSSDGSWKTNFL 240
Db 181 EWTLLTYKHQTKDYTKCVLFGNMIAYFLSKKICTSPRGGYFLSSDGSWKTNFL 240
Qy 241 KEGERHLVSKLYTDDMPETVTTTAQETKGRIOTKKEVAIKTTLKELVHKRVTSPE 300
Db 241 KEGERHLVSKLYTDEMPEVTTTVAQAKGRIOTKRKEVSIKTLKELVHKRVTSPE 300
Qy 301 DWMNQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSKLTNFSLPD 360
Db 301 DWMNQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSKLTNFSLPD 360
Qy 361 TRTCRIFAFHGWNVYKCHAI CCVLNRQGGKRNVLPHGPASTGKSIIAQIAQAVGNVG 420
Db 361 TRTCRIFAEHGWNVYKCHAI CCVLNRQGGKRNVLPHGPASTGKSIIAQIAQAVGNVG 420
Qy 421 CYNAAANVFPNDCTNKNLIWVEEAGNFQGVNQFKAICSGQTIIRDQKKGSKQIEPTP 480
Db 421 CYNAAANVFPNDCTNKNLIWVEEAGNFQGVNQFKAICSGQTIIRDQKKGSKQIEPTP 480
```

```

QY 481 VMTNNTVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVK 540
DB 481 VMTNNTVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVK 540
QY 541 NGYSTMASYCAKMGKVPDWSNWAEPKVPPTINLLGSARSPTTPKSTPLSQNYALTPL 600
DB 541 NGYSTMACYCAKMGKVPDWSNWAEPKLETPINSLSMRSPSLTPRSTPLSQNYALTPL 600
QY 601 ASDLELALPEWSTPNTPVAGTAETQNTGAGSKACQDQGLSTWSEIEDLRACFGABP 660
DB 601 ASDLADLALPEWSTPNTPVAGTAASQNTGAGSTACQAGRSPTWSEIEDLRACFSQEQ 660
QY 661 LKDFSEPLNLD 672
DB 661 LKDFSDSLTLD 672

RESULT 10
VNC5_PAVH STANDARD; PRT; 672 AA.
AC P03133;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVP1).
GN NamesNS1;
OS Hamster parvovirus H1.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10799;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83112183; PubMed=6823009;
RA Rhode S.L. III, Paradiso P.R.;
RT "Parvovirus genome: nucleotide sequence of H-1 and mapping of its
RT genes by hybrid-arrested translation.";
RL J. Virol. 45:173-184 (1983).
CC -!- FUNCTION: Seems necessary for viral DNA replication.
CC -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; X01457; CA925689.1; --
DR PIR; A03695; UTPV1.
DR InterPro; IPR001257; Parvo_NS1.
DR Pfam; PF01057; Parvo_NS1; 1.
KW ATP-binding; DNA replication; Noncapsid protein;
KW Nonstructural protein.
FT NP_BIND 399 406 ATP (Potential).
SQ SEQUENCE 672 AA; 75993 MW; 12F331142F72AA6D CRC64;

Query Match 91.5%; Score 3328; DB 1; Length 672;
Best Local Similarity 91.1%; Pred. No. 1.3e-238;
Matches 612; Conservative 25; Mismatches 35; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVLGATNWLKESNQEVFSFVFPKENVQLNGKIDGWSYKKEQLDELKSLQ 60
DB 1 MAGNAYSDEVLGVTNMLKDKSQSEVFSFVFPKENVQLNGKIDGWSYKKEQLDELKSLQ 60
QY 61 RGAETTWDSQEDMEWETTVDMTKQVIFDLSLVKKCLFVLATKNIIFGCDVNVWFVQHEW 120
DB 61 RGAETTWDSQEDMEWESAVDMTKQVIFDLSLVKKCLFVLSTKNIAPSNVTFVQHEW 120
QY 121 GKQDGHCHVLIGKDFPSQAGKWRRLQNLVYNSRWLVACNVQLTPAERIKLREIAEDN 180
DB 121 GKDPGHCHVLIGKDFPSQAGKWRRLQNLVYNSRWLVACNVQLTPAERIKLREIAEDS 180
QY 181 EWTLLTYKHKTKDYTKCVLFGNMIAYYFLTKKIKSTSPPRDGGYFLSSDGSWKTNFL 240

```

```

DB 181 EWTLLTYKHKTKDYTKCVLFGNMIAYYFLSKKIKSTSPPRDGGYFLSSDGSWKTNFL 240
QY 241 KEGERHLVSKLYTDDMRPETVTTTAQETKRGRIQTKKEVAIKTTLKELVHKRVTSPE 300
DB 241 KEGERHLVSKLYTDEMKEPETVTTTAQEAQKRGRIQTRREVSIKTTLKELVHKRVTSPE 300
QY 301 DWMMPQDSYIEWMAQPGGENLLKNTLEICTLTARTKTAFDLILKAETSKLTNFSLPD 360
DB 301 DWMMPQDSYIEWMAQPGGENLLKNTLEICTLTARTKTAFDLILKAETSKLANFSMAS 360
QY 361 TTRCIRFAFHGNNYKVCCHAI CCVLNRQGGKNTVLFHGPASTGKSI IAQAIAQAVGNVG 420
DB 361 TTRCIRFAEHGNNYIKVCHAI CCVLNRQGGKNTVLFHGPASTGKSI IAQAIAQAVGNVG 420
QY 421 CYNAAANVFPNDCTNKNLIWVEEAGNFQGVQVQNFKAICSGQTIRIDQKKGSKQIEPTP 480
DB 421 CYNAAANVFPNDCTNKNLIWVEEAGNFQGVQVQNFKAICSGQTIRIDQKKGSKQIEPTP 480
QY 481 VMTNNTVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVK 540
DB 481 VMTNNTVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVK 540
QY 541 NGYSTMASYCAKMGKVPDWSNWAEPKVPPTINLLGSARSPTTPKSTPLSQNYALTPL 600
DB 541 NGYSTMACYCAKMGKVPDWSNWAEPKLDTPINSLSMRSPSLTPRSTPLSQNYALTPL 600
QY 601 ASDLELALPEWSTPNTPVAGTAETQNTGAGSKACQDQGLSTWSEIEDLRACFGABP 660
DB 601 ASDLADLALPEWSTPNTPVAGTAASQNTGAGSTACQAGRSPTWSEIEDLRACFSQEQ 660
QY 661 LKDFSEPLNLD 672
DB 661 LKDFSEELTLD 672

RESULT 11
Q8JUV16 PRELIMINARY; PRT; 672 AA.
ID Q8JUV16;
AC 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nonstructural protein 1.
GN NamesNS1;
OS Rat minute virus 1b.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=172386;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22120170; PubMed=12124471;
RA Wan C.H., Soderlund-Venermo M., Pintel D.J., Riley L.K.;
RT "Molecular characterization of three newly recognized rat
RT parvoviruses.";
RL J. Gen. Virol. 83:2075-2083 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Wan C.-H., Pintel D.J., Soderlund-Venermo M., Riley L.K.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF332883; AM93277.1; --
DR GO; GO:0019012; C:virion; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR001257; Parvo_NS1.
DR Pfam; PF01057; Parvo_NS1; 1.
KW Nonstructural protein.
SQ SEQUENCE 672 AA; 76201 MW; C2F1A71F6BEF449A6 CRC64;

Query Match 91.4%; Score 3325; DB 2; Length 672;
Best Local Similarity 90.9%; Pred. No. 2.1e-238;
Matches 611; Conservative 25; Mismatches 36; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVLGATNWLKESNQEVFSFVFPKENVQLNGKIDGWSYKKEQLDELKSLQ 60

```

Db 1 MAGNAYSDEVLGATNWLKDKSQSEVFSFVFKENNVQLNGKDIGWNSYRKELQDELKSLQ 60
Qy 61 RGAETTWDSQSDMEWETTVDWTKKQVFIPLSLVKKCLFEVLNTKNIPLFGDVNVFVQHEW 120
Db 61 RGAETTWDSQSDMEWESAVDDMTKKQVFIPLSLVKKCLFEVLNTKNIPLFSDVTVFVQHEW 120
Qy 121 GKDGQWHCHVLIGGKDFSAQGWRRQLNVYWSRWLVATCNVQLTPAERIKLREIAEDN 180
Db 121 GKDGQWHCHVLIGGKDFSAQGWRRQLNVYWSRWLVATCNVQLTPAERIKLREIAEDS 180
Qy 181 EWTLLTYKHQTKDYKCVLFGNMIAYFLTKKISTSPRGGYFLSSDSGKNTNPL 240
Db 181 EWTLLTYKHQTKDYKCVLFGNMIAYFLTKKISTSPRGGYFLSSDSGKNTNPL 240
Qy 241 KEGERHLVSKLYTDDMRPETVETVTTAQTGKRIOTKKEVALKTLKELVHKRVTSPE 300
Db 241 KEGERHLVSKLYTDDMRPETVETVTTAQTGKRIOTKKEVALKTLKELVHKRVTSPE 300
Qy 301 DMMWQPDYSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAEYSKLTNPLSD 360
Db 301 DMMWQPDYSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAEYSKLTNPLSD 360
Qy 361 TRTCRIFAFAHGWNYKVCCHAI CCVLNRQGGKNTVLFHGPASTGKSIIAQAIAQAVNGV 420
Db 361 TRTCRIFAFAHGWNYKVCCHAI CCVLNRQGGKNTVLFHGPASTGKSIIAQAIAQAVNGV 420
Qy 421 CYNAAVNFPPNDCTNKNLIWBEAGNFGQVNOFKALCSGQTIIDOKGKSGKQIEPTP 480
Db 421 CYNAAVNFPPNDCTNKNLIWBEAGNFGQVNOFKALCSGQTIIDOKGKSGKQIEPTP 480
Qy 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
Db 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
Qy 541 NGYQSTWASCYCAKWKVPDSEWAEKVPVTPINLGSARSPTTPKSTPLSQNYALTPL 600
Db 541 NGYQSTWASCYCAKWKVPDSEWAEKVPVTPINLGSARSPTTPKSTPLSQNYALTPL 600
Qy 601 ASDLEDLALPWSPTNPVAGTAEONTGAGSKACODGQLSPTWSEIEEDLRACFGAEP 660
Db 601 ASDLEDLALPWSPTNPVAGTAEONTGAGSKACODGQLSPTWSEIEEDLRACFGAEP 660
Qy 661 LKPDFSEPLNLD 672
Db 661 LKPDFSEPLNLD 672

RESULT 12

VNCS PAVL3 STANDARD; PRT; 668 AA.
AC P36311;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1).
GN Name=NS1;
OS Parvovirus LuIII.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=35339;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93297126; PubMed=8517025;
RA Difford N., Chen K.C., Bates R.C., Lederma M.;
RT "The complete nucleotide sequence of parvovirus LuIII and localization
of a unique sequence possibly responsible for its encapsidation
pattern.";
RT Virology 192:339-345(1993).
RL -/- FUNCTION: Seems necessary for viral DNA replication.
CC -/- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC -/- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.emb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----

DR EMBL; M81888; -; NOT ANNOTATED_CDS.

DR PIR; A44276; A44276.

DR InterPro; IPR001257; Parvo_NSI.

DR Pfam; PF01057; Parvo_NSI; 1.

KW ATP-binding; DNA replication; Noncapsid protein;

KW Nonstructural protein.

FT NP BIND 399 406 ATP (Potential).

SQ SEQUENCE 668 AA; 75846 MW; CAB69049F8F86B53 CRC64;

Query Match 91.4%; Score 3324; DB 1; Length 668;

Best Local Similarity 91.4%; Pred. NO. 2.5e-238;

Matches 614; Conservative 24; Mismatches 30; Indels 4; Gaps 2;

Qy 1 MAGNAYSDEVLGATNWLKDKSQSEVFSFVFKENNVQLNGKDIGWNSYRKELQDELKSLQ 60

Db 1 MAGNAYSDEVLGATNWLKDKSQSEVFSFVFKENNVQLNGKDIGWNSYRKELQDELKSLQ 60

Qy 61 RGAETTWDSQSDMEWETTVDWTKKQVFIPLSLVKKCLFEVLNTKNIPLFGDVNVFVQHEW 120

Db 61 RGAETTWDSQSDMEWESSVDELTKKQVFIPLSLVKKCLFEVLNTKNIPLFGDVNVFVQHEW 120

Qy 121 GKDGQWHCHVLIGGKDFSAQGWRRQLNVYWSRWLVATCNVQLTPAERIKLREIAEDN 180

Db 121 GKDGQWHCHVLIGGKDFSAQGWRRQLNVYWSRWLVATCNVQLTPAERIKLREIAEDQ 180

Qy 181 EWTLLTYKHQTKDYKCVLFGNMIAYFLTKKISTSPRGGYFLSSDSGKNTNPL 240

Db 181 EWTLLTYKHQTKDYKCVLFGNMIAYFLTKKISTSPRGGYFLSSDSGKNTNPL 240

Qy 241 KEGERHLVSKLYTDDMRPETVETVTTAQTGKRIOTKKEVALKTLKELVHKRVTSPE 300

Db 241 KEGERHLVSKLYTDDMRPETVETVTTAQTGKRIOTKKEVALKTLKELVHKRVTSPE 300

Qy 301 DMMWQPDYSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAEYSKLTNPLSD 360

Db 301 DMMWQPDYSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAEYSKLTNPLSD 360

Qy 361 TRTCRIFAFAHGWNYKVCCHAI CCVLNRQGGKNTVLFHGPASTGKSIIAQAIAQAVNGV 420

Db 361 TRTCRIFAFAHGWNYKVCCHAI CCVLNRQGGKNTVLFHGPASTGKSIIAQAIAQAVNGV 420

Qy 421 CYNAAVNFPPNDCTNKNLIWBEAGNFGQVNOFKALCSGQTIIDOKGKSGKQIEPTP 480

Db 421 CYNAAVNFPPNDCTNKNLIWBEAGNFGQVNOFKALCSGQTIIDOKGKSGKQIEPTP 480

Qy 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540

Db 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540

Qy 541 NGYQSTWASCYCAKWKVPDSEWAEKVPVTPINLGSARSPTTPKSTPLSQNYALTPL 600

Db 541 NGYQSTWASCYCAKWKVPDSEWAEKVPVTPINLGSARSPTTPKSTPLSQNYALTPL 600

Qy 601 ASDLEDLALPWSPTNPVAGTAEONTGAGSKACODGQLSPTWSEIEEDLRACFGAEP 660

Db 598 -SDLEDLALPWSPTNPVAGTAEONTGAGSKACODGQLSPTWSEIEEDLRACFGAEP 660

Qy 661 LKPDFSEPLNLD 672

Db 657 WKSDESEPLNLD 668

RESULT 13

O71159

ID O71159 PRELIMINARY; PRT; 665 AA.

AC O71159;

DT 01-AUG-1998 (TrEMBLrel. 07, Created)

DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)

DT	01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE	Nonstructural protein (fragment).
GN	Name=NS1;
OS	Kilham rat virus.
OC	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX	NCBI_TaxID=12441;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=U-Mass;
RX	MEDLINE=98184569; PubMed=9525656;
DR	Ball-Goodrich L.J., Leland S.E., Johnson E.A., Paturzo F.X.,
RA	Jacoby R.O.;
RT	"Rat parvovirus type 1: the prototype for a new rodent parvovirus
RT	serogroup.";
RL	J. Virol. 72:3289-3299(1998).
RL	EMBL; AF036711; AAC40695.1; -.
DR	GO: GO:0019012; C:virion; IEA.
DR	GO: GO:0019079; P: viral genome replication; IEA.
DR	InterPro; IPR001257; Parvo_NSR.
DR	Pfam; PF01057; Parvo_NSR; 1.
KW	Nonstructural protein.
FT	NON TER
FT	1
SO	SEQUENCE 665 AA; 75375 MW; 778E29043417E409 CRC64;
Query Match 90.8%; Score 3305; DB 2; Length 665;	
Best Local Similarity 91.4%; Pred. No. 6.4e-237;	
Matches 608; Conservative 23; Mismatches 34; Indels 0; Gaps 0;	
Qy	8 DEVLGATNWLKEKSNQVFSVFVKNNVQLNGKDIGNWSYKKEIQEDLKSLQGAETTW 67
Db	1 DEVLGATNWLKDKSQVFSVFVKNNVQLNGKDIGNWSYKKEIQEDLKSLQGAETTW 60
Qy	68 DQSEDMEWETTVDSEMTKKQVFIQFDSLVKKCLFEVLNTKNIIPGDVNVFVQHEWGKDQGW 127
Db	61 DQSEDMEWESAVDDMTKKQVFIQFDSLVKKCLFEVLNTKNIIPGDVNVFVQHEWGKDQGW 120
Qy	128 CHVLIGKGFSAQAGKWRROLNVYSGRWLTACNVQLTPAERIKLRIABDNSEWVTLT 187
Db	121 CHVLIGKGFSAQAGKWRROLNVYSGRWLTACNVQLTPAERIKLRIABDNSEWVTLT 180
Qy	188 YKHQTKKDYTKCVLFGNMIAIYELTKKKISTSPRGGYFLSSDGSWKNTFLKEGRHL 247
Db	181 YKHQTKKDYTKCVLFGNMIAIYELTKKKISTSPRGGYFLSSDGSWKNTFLKEGRHL 240
Qy	248 VSKLYTDDMRPETVETTTVAQETKRGRIQTKKEVAIKTTLKELVHKRVTSPEWMMWQP 307
Db	241 VSKLYTDDMRPETVETTTVAQETKRGRIQTKKEVAIKTTLKELVHKRVTSPEWMMWQP 300
Qy	308 DSYIEMMAQPGENLLKNTLBEICTLTARTKTAFDLILEKAETSKLNFSLPDRTRCRIF 367
Db	301 DSYIEMMAQPGENLLKNTLBEICTLTARTKTAFDLILEKAETSKLNFSLPDRTRCRIF 360
Qy	368 AFHGMNVKVCCHATCCVLNRGGKRVTLFHGPASTGKSIITAOIAQAGVNGVCYNAANV 427
Db	361 AEHGMNVKVCCHATCCVLNRGGKRVTLFHGPASTGKSIITAOIAQAGVNGVCYNAANV 420
Qy	428 NFPNDCTNKNLIWVEAGNFGQVNOFPKALCSGOTIRIDQKGSQKIEPTPVMITNE 487
Db	421 NFPNDCTNKNLIWVEAGNFGQVNOFPKALCSGOTIRIDQKGSQKIEPTPVMITNE 480
Qy	488 NITVVRIGCEERPEHTQPIRDRMLNIHLTHHLPGDGLVDKNEWPMICAWLVKNGYOSTM 547
Db	481 NITVVRIGCEERPEHTQPIRDRMLNIHLTHHLPGDGLVDKNEWPMICAWLVKNGYOSTM 540
Qy	548 ASYCAKKGKVPDWSENWAEKPVPTPINLLGSARSFPTPKSTPLSONYALTPLASDL 607
Db	541 ASYCAKKGKVPDWSEDWAEKPLETPIINSLGSMRSPSLTPRSTPLSONYALTPLASDL 600
Qy	608 ALEPWSPTNPVAGTAEQNTGAEAGSKACODQLSPTWSETEEDLRACFGEAPLKKDPSE 667
Db	601 ALEPWSPTNPVAGTAEQNTGAEAGSTACQGAQRSPTWSETEADLRACFSQEQLESDFNE 660
Qy	668 PLNLD 672

```

Db          661 661 ELTLD 665  || ||
RESULT 14
071157      PRELIMINARY;      PRT;      671 AA.
ID 071157;
AC 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Nonstructural protein.
GN NamesNS1;
OS Rat parvovirus la.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=74581;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98184569; PubMed=9525656;
RA Ball-Goodrich L.J., Leland S.B., Johnson E.A., Paturzo F.X.,
RA Jacoby R.O.;
RT "Rat parvovirus type 1: the prototype for a new rodent parvovirus
RT serogroup";
RL J. Virol. 72:3289-3299 (1998).
DR EMBL; AF036710; AAC40693.1; -.
DR GO; GO:0019012; C:virion; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo_NS1; I.
DR Nonstructural protein.
KW
SQ SEQUENCE 671 AA; 75752 MW; 9BCB39A39298D4DE CRC64;

Query Match      80.7%; Score 2937.5; DB 2; Length 671;
Best Local Similarity 81.1%; Pred. No. 1.4e-209;
Matches 545; Conservative 43; Mismatches 83; Indels 1; G

Qy 1 MAGNAYSDEVILGATNWLKSKNSQEVFVFPKNENVQLNGKDTGWSYKKELOEDL
Db 1 MAGNQYSEBVLIDGVNWLKQDSQTEAFSVFKCNVQLNGKIDINWNNYNKPLNNEEL
Qy 61 RGAETTWQSDSEDMKEWTTVDMTKKQVIFDSLVKKCLFEVLNTKNIIPGDVNVFV
Db 61 RGAETAMQDAEDMEWSELDTLTKKQVIFDALVKKCLFEVLINKNVAPNDCGWFI
Qy 121 GKDQGWCHVLITGGKDFPSQAQGWRRQLNVYMSRLVLTACNVQLTTPAERIKLREI
Db 121 GKDQGWCHVLHGKGLQAGKWFRRQMNVLMSRLVLTACSVNLSPAERIKLREI
Qy 181 EWVTLTYKHQTKKDYTKCVLFGNMIAYYFLFKKIKISTSPRDDGYSFLSSDSGKW
Db 181 EWVTLTYRHKQTKKDYVKLVHFGNMIAYYFMTKKIIASPPRDGYSFLSTDGSKW
Qy 241 KEGEHLVSKLYTDDMPETVETVTTAOSTKGRLOTQKEVAIKTTLKELVHKVE
Db 241 KEADRHVLSLYTDQMKPETVETVTTAQSTKGRLOTQKEVSIKCTLRDLVAKRV
Qy 301 DNMWQPPDSYIEMWAPQGENLLKNTLEICTLTLARTKTAFDLILEKAETSKLTNFP
Db 301 DNMWQPPDSYIEMWAPQGENLLKNTLEICTLTLARTKTAFDLILEKAETSKLTTF
Qy 361 TTRTCRIFAFHGNVYKVCHAIACCVLNRQGGKRNVLFHGPASTGKSIIAQAIQAV
Db 361 TTRTCRIFAGHGWNVYKCHAIACVLNRQGGKRNVLFHGPASTGKSIIAQAIQAV
Qy 421 CYNAANVFPNDCTNKNLWVEEAGNFGQVNOQFKAICSGOTIRIDQKGKSKOI
Db 421 CYNAANVFPNDCTNKNLWVEEAGNFGQVNOQFKAICSGOTIRIDQKGKSKOI
Qy 481 VIMTTNENITVVRIGCSERBEHTQPIRDRMLNTHLTHLPDGFGLVDKNEMPMICA
Db 481 VIMTTNENITVVRIGCSERBEHTQPIRDRMLNHLNTHLPDGFGLVEKQEPFLICA
Qy 541 NGYQSTMASYCAKMGKVPDWSENWAEKVPFTPINLIGSARSFFTPKSTPLSQNTA

```

```
Db 541 NGYQSTWASYSHHWGKVPDMSENWABEPAMOTPINISGARSTQSTATSTPLSQNYALTPL 600
Qy 601 ASDLEDLALBPWSTPNTPVAGTAETQNTGEAGSKACODGQLSPTWSEIEDLRACFGAEP 660
Db 601 ASDLADLALBPWSTPNTPVAEPTVGSNT-NTGGRNSQTARASPTWSEIADLRACFSQEQ 659
Qy 661 LKQDFSEPLNLD 672
Db 660 LESDFNEELTLD 671

RESULT 15
ID P89513 PRELIMINARY; PRT; 668 AA.
AC P89513;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nonstructural protein 1.
OS Feline panleukopenia virus (FPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10786;
RN [1]
RP SEQUENCE FROM N.A.
RA Horiuchi M.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB000053; BAA19014.1; -.
DR GO; GO:0019012; C:virion; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0019079; P:virial genome replication; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; I.
DR SMART; SM00382; AAA_1.
DR Nonstructural protein.
KW SEQUENCE 668 AA; 76743 MW; 84A2A10104D9D5788 CRC64;
SQ
```

```
Query Match 73.0%; Score 2657; DB 2; Length 668;
Best Local Similarity 73.4%; Pred. No. 1e-188;
Matches 493; Conservative 71; Mismatches 102; Indels 6; Gaps 5;

Qy 1 MAGNAYSDEVILGATNWLKESNOEVSFVPKENVOLNGKDIGNSYKKELOEDELKSLQ 60
Db 1 MSGNQYTEEVMEGVNWLKKAENAEAFSFVKCDNVOLNGKDVRRNNYTKPIQNEELTSLI 60

Qy 61 RGAETTWDS--EDMEWETTVDGMTKQVFIPLSLVKCLFEVLNTKNIPPGDVNWFVCH 118
Db 61 RGAQTAMDQTEEMDEWESEVDSLAKQVQTFDALIKKCLFEVFSKNIEPNECVWFIOH 120

Qy 119 EMGKQGWCHVLIGGKFSQAQKWRRLQNVYNSRWLVATCNVOLTTPAERIKLREIAE 178
Db 121 EMGKQGWCHVLLHLSKNLQATGKWLRRQNMWYSRWLVTLCSVNLTPTEKIKLREIAE 180

Qy 179 DNEWVLLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPDRDGGYFLSSDSGWKN 238
Db 181 DSEWVTILYRHKQTKDYKVMVHFGNMIAYFLTKKISTSPDRDGGYFLSSDSGWKN 239

Qy 239 FLKGEGERHLVSKLYTDDMPETVETVTTAQTCKGRIOTKKEVAIKTTLKELVHKRVTS 298
Db 240 FMKHQDRHVSTLYTEQMPEVETVTTAQTCKGRIOTKKEVSIKCTLRDLVSKRVTS 299

Qy 299 PEDMMMQPDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAEYSKLTNFSL 358
Db 300 PEDMMMLQPDYIEMMAQPGGENLLKNTLEICTLTARTKTAFELILEKADNTKLTNFDL 359

Qy 359 PDTRTCRIEAFHGNVYKVCIAICCVLRQGGKRNVLPHGPASTGKSIIAQIAQVGN 418
Db 360 ANSRTCQIFPMFGWNWIKVCIAIACVLRQGGKRNVLPHGPASTGKSIIAQIAQVGN 419

Qy 419 VGCYNAANVNFENDCTNKNLIWEAGNFGQVQVQFKAICSGQITRIDOKGKSKQIEP 478
Db 420 VGCYNAANVNFENDCTNKNLIWEAGNFGQVQVQFKAICSGQITRIDOKGKSKQIEP 479
```

```
Qy 479 TPVIMTTNENITVVRIGCERPEHTOPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWL 538
Db 480 TPVIMTTNENITVVRIGCERPEHTOPIRDRMLNIHLVCKLPGDFGLVDKEEWPICAWL 539
Qy 539 VKNGYQSTWASYCAKWKVPDMSENWABEPKVPPTPINLLGSARSFPFTTPKSTPLSQNYALT 598
Db 540 VKHGYQSTWANYTHHWGKVPEDENWAEPKIQEGINSPG-CXDLETQAASNPOSQDHVLT 598
Qy 599 PLASDLEDLALBPWSTPNTPVAGTAETQNTGEAGSKACODGQLSPTWSEIEDLRACFGA 658
Db 599 PLTPDVVDLALBPWSTPNTPTPIAETA-NOQSNQLG-VTHKDVQASPTWSEIADLRIFTS 656
Qy 659 EPLKKDFSEPLN 670
Db 657 EQLEEDFRDLD 668
```

Search completed: January 22, 2005, 03:40:43
Job time : 110.333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2005, 23:46:57 ; Search time 8541.75 Seconds
(without alignments)
11177.792 Million cell updates/sec

Title: US-10-069-056-8
Perfect score: 2019
Sequence: 1 atggctggaatgcttactc.....agccgctgaacttgactaa 2019

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	2019	100.0	2019	6	AX137743 Sequence
2	2017.4	99.9	2019	6	AX137736 Sequence
3	2017.4	99.9	5081	14	VO1115 Minute viru
4	2017.4	99.9	5149	14	J02275 Minute viru
5	2015.8	99.8	2019	6	AX137739 Sequence
6	2015.8	99.8	2019	6	AX137747 Sequence
7	2015.8	99.8	2019	6	AX137751 Sequence
8	1905.4	94.4	5085	14	M12032 Minute viru
9	1902.2	94.2	4764	14	M0U34253
10	1902.2	94.2	5087	14	PAMVMI
11	1899	94.1	5144	14	MPUI2469
12	1873	92.8	4764	14	M0U34254
13	1862.2	92.2	4773	14	H0U34255
14	1855.8	91.9	4761	14	MVU34256
15	1771	87.7	5121	6	CQ786765
16	1662.2	82.3	4904	14	AF321230
17	1659	82.2	4795	14	AF321284
18	1657.4	82.1	4927	14	KR079033
19	1656.6	82.1	5135	14	PVRSEQ

20	1655.8	82.0	4816	14	AF332883	AF332883 Rat minut
21	1652.6	81.9	4813	14	AF332882	AF332882 Rat minut
22	1644.6	81.5	5176	14	PARHI	X01457 Parvovirus
23	1644.4	81.4	3995	14	AF036711	AF036711 Kilham ra
24	1275	63.2	4936	14	AF036710	AF036710 Rat parvo
25	1051.6	52.1	2007	14	AB000062	AB000062 Feline pa
26	1050	52.0	2007	14	AB000048	AB000048 Feline pa
27	1050	52.0	2007	14	AB000057	AB000057 Feline pa
28	1050	52.0	5124	14	PVPVC	M38246 Feline panl
29	1046.8	51.8	2007	14	AB000049	AB000049 Feline pa
30	1046.8	51.8	2007	14	AB000058	AB000058 Feline pa
31	1046.8	51.8	2007	14	AB000060	AB000060 Feline pa
32	1046.8	51.8	2007	14	AB000069	AB000069 Feline pa
33	1045.2	51.8	2007	14	AB000053	AB000053 Feline pa
34	1045.2	51.8	2007	14	AB000063	AB000063 Feline pa
35	1045.2	51.8	2007	14	AB000065	AB000065 Feline pa
36	1045.2	51.8	2007	14	AB000067	AB000067 Feline pa
37	1045.2	51.8	5075	14	PVCYIA	D26079 Canine parv
38	1043.6	51.7	2007	14	AB000051	AB000051 Feline pa
39	1043.6	51.7	4983	14	PAPVNS1	X55115 Feline panl
40	1043.6	51.7	5124	14	PVPVC	M38245 Canine parv
41	1042	51.6	2007	14	AB000055	AB000055 Feline pa
42	1042	51.6	5323	14	PVCCPN	M19296 Canine parv
43	1040.4	51.5	5049	6	AR043629	AR043629 Sequence
44	1040.4	51.5	5049	6	AR043630	AR043630 Sequence
45	1040.4	51.5	5094	14	PVRFD	D00765 Mink enteri

ALIGNMENTS

RESULT 1	AX137743	Sequence 8 from Patent EP1077260.	2019 bp	DNA	linear	PAT 30-MAY-2001
LOCUS	AX137743					
DEFINITION	AX137743					
ACCESSION	AX137743.1	GI:14273917				
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						

Mice minute virus
Mice minute virus
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.

REFERENCE 1
Nuesch, J. and Rommelaere, J.
Parvovirus ns1 variants
Patent: EP 1077260-A 8 21-FEB-2001;
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts
(DE)

FEATURES
source

Location/Qualifiers
1..2019
/organism="Mice minute virus"
/mol_type="unassigned DNA"
/db_xref="taxon:10794"
/note="Parvovirus NS1 variant"
1..2019
/note="unnamed protein product"

CDS

/protein_id="CAC39993.1"
/db_xref="GI:14273918"
/translation="MAGNAYSDEVLGATNWLKEKSNQEVFSFVKENVOLNGKDIGW
NSYKKELODELKSLQRGAEETWQSEDMETTTVDGMTKKQVFIQDSLVKKCLFEVL
NTNIFPGDVNMFVQHWKQDQGHCHVLIGKDFQAQGWRRQLNVTWRSWLVA
CNVQLTPAERIKLREIAEDNEWTLTYKHQTKDYTKCVLFGNMIAIYFLTKKKIS
TSPRDRGGYFLSDSGWNTFKGERHLVSKLYTDDMRPTEVTTVTAQETKGRIS
QTKKEVSIKTLKELVHKVTSPEDEMMQPSYIEMMAQPGGKNLLKNTLETCTLT
ARTKAFDILILEKETS KLNTSLPDRACRIFAFHGMVYKVCHAI CCVLNRQGGKR
NTVLEHGPASTGKSIIAQIAQAVGNVCYNAANVFNENDCTNKLIWVEEAGNFQ
QVQFKACSGQITRIDQKSGKQIEPTVIMTNTNENTVVRIGKEERPEHTQPIRD
RLNLTHTLTPDGLVGNWPMI CAWLXNGYQSTWASYCAKMGKVPDWTENNAE
PKVPTPINLIGARSPTFKSTPLSQNTALPTLASDLEDLALPELSTNTVPAGTAE
TQNTGSGSKACQDQGLSPTWSEIEEDLRACFGAEPLKKDFSEPLND"

ORIGIN

Query Match 100.0%; Score 2019; DB 6; Length 2019;

Best Local Similarity 100.0%; Pred. No. 0;			
Matches 2019; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	ATGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA	60
Db	1	ATGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA	60
Qy	61	AGTAACGAGGAAGTGTCTCATTTGTTTTTAAATAAGAAATGTTCAACTGGAATGGA	120
Db	61	AGTAACGAGGAAGTGTCTCATTTGTTTTTAAATAAGAAATGTTCAACTGGAATGGA	120
Qy	121	GATATCGGATGGAATAGTTTACAAAAGAGCTGAGGAGCAAGCTGGAATCTTTACAA	180
Db	121	GATATCGGATGGAATAGTTTACAAAAGAGCTGAGGAGCAAGCTGGAATCTTTACAA	180
Qy	181	CGAGGAGCGGAATCTCTCGGACCAAGAGGAGCAATGGAATGGGAAACCAAGTGGAT	240
Db	181	CGAGGAGCGGAATCTCTCGGACCAAGAGGAGCAATGGAATGGGAAACCAAGTGGAT	240
Qy	241	GAATGACCAAAAAGCAAGTAATTTTCAATTTTGGTTTAAATAAGTATTTTAA	300
Db	241	GAATGACCAAAAAGCAAGTAATTTTCAATTTTGGTTTAAATAAGTATTTTAA	300
Qy	301	GTGCTTAAACAAGAAATATTTCTGCTGATGTTAAATTTGTTTGTGCAACATGAATGG	360
Db	301	GTGCTTAAACAAGAAATATTTCTGCTGATGTTAAATTTGTTTGTGCAACATGAATGG	360
Qy	361	GGAAAGACCAAGCTGGCACTGCCATGTACTTAATTTGGAGGAAGCACTTTAGTCAAGCT	420
Db	361	GGAAAGACCAAGCTGGCACTGCCATGTACTTAATTTGGAGGAAGCACTTTAGTCAAGCT	420
Qy	421	CAAGGGAATGTTGGAGAGGCACTAAATTTGTTTACTGGAGCAGATGGTTGTAACAGCC	480
Db	421	CAAGGGAATGTTGGAGAGGCACTAAATTTGTTTACTGGAGCAGATGGTTGTAACAGCC	480
Qy	481	TGTAATGTGCACTAACACAGCTGAAGAAATTAACCTAAGAAATAGCAGAAGCAAT	540
Db	481	TGTAATGTGCACTAACACAGCTGAAGAAATTAACCTAAGAAATAGCAGAAGCAAT	540
Qy	541	GAGTGGTACTCTACTTACTATAGCATAAGCAACCAAAAGAACTATACCAAGTGT	600
Db	541	GAGTGGTACTCTACTTACTATAGCATAAGCAACCAAAAGAACTATACCAAGTGT	600
Qy	601	GTTCCTTTTGGAAACATGATGCTTACTATTTTAACTAAAGAAATAAGCACTAGT	660
Db	601	GTTCCTTTTGGAAACATGATGCTTACTATTTTAACTAAAGAAATAAGCACTAGT	660
Qy	661	CCACCAAGAGACGGAGCTATTTCTTAGCAGTACTCTGGCTGGAAAACTAACTTTTA	720
Db	661	CCACCAAGAGACGGAGCTATTTCTTAGCAGTACTCTGGCTGGAAAACTAACTTTTA	720
Qy	721	AAAGAAAGGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG	780
Db	721	AAAGAAAGGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG	780
Qy	781	GTTGAAACCACTAACCCTGCGAGGAACTAAGCGGGCAGAAATTCAAAATAAAAAA	840
Db	781	GTTGAAACCACTAACCCTGCGAGGAACTAAGCGGGCAGAAATTCAAAATAAAAAA	840
Qy	841	GAAGTTTCTATTAAACTACATTTAAGAGCTGGTGCAATAAAGATTAACCTCACAGAG	900
Db	841	GAAGTTTCTATTAAACTACATTTAAGAGCTGGTGCAATAAAGATTAACCTCACAGAG	900
Qy	901	GACTGGATGATGATGAGCGAGACAGTTTCAATTTGAAATGATGCTCAACAGGTGGAGAA	960
Db	901	GACTGGATGATGATGAGCGAGACAGTTTCAATTTGAAATGATGCTCAACAGGTGGAGAA	960
Qy	961	AACCTGCTGAAAAATACGCTAGAGATTTGTACATACTAAGCTAGGCAAGCAACCAAGCA	1020
Db	961	AACCTGCTGAAAAATACGCTAGAGATTTGTACATACTAAGCTAGGCAAGCAACCAAGCA	1020
Qy	1021	TTTGACTTAATTTTAAAAAGCTGAAACCAAGCAACCACTAACCACTTTTCACTGCTGAC	1080
Db	1021	TTTGACTTAATTTTAAAAAGCTGAAACCAAGCAACCACTAACCACTTTTCACTGCTGAC	1080

Db	1021	TTTGACTTAATTTTAAAAAGCTGAAACCAAGCAACCACTAACCACTTTTCACTGCTGAC	1080
Qy	1081	ACAAGAGCTGCAAGAAATTTTGTCTTTTCAATGCTGGAATCTATGCTAAAGTTTCCCATGCT	1140
Db	1081	ACAAGAGCTGCAAGAAATTTTGTCTTTTCAATGCTGGAATCTATGCTAAAGTTTCCCATGCT	1140
Qy	1141	ATTGCTGTGTTTTTAAACAGACAGGAGGCAAAAGAAATCTGTTTTTATTTCACTGAGCA	1200
Db	1141	ATTGCTGTGTTTTTAAACAGACAGGAGGCAAAAGAAATCTGTTTTTATTTCACTGAGCA	1200
Qy	1201	GCAGACAGGCAAGTAATTTATTTGCACAAGCCATAGCAAGAGCTGGCAATGTTGGT	1260
Db	1201	GCAGACAGGCAAGTAATTTATTTGCACAAGCCATAGCAAGAGCTGGCAATGTTGGT	1260
Qy	1261	TGCTATAATGCGAGCAATGTAACTTTTCCATTTTAAATGCTGTACCAAGAACTTGATT	1320
Db	1261	TGCTATAATGCGAGCAATGTAACTTTTCCATTTTAAATGCTGTACCAAGAACTTGATT	1320
Qy	1321	TGGGTAGAAAGAGCTGTTAACTTTTGGACAGCAAGTAAGCACTTTAAAGCCATTTGCTCT	1380
Db	1321	TGGGTAGAAAGAGCTGTTAACTTTTGGACAGCAAGTAAGCACTTTAAAGCCATTTGCTCT	1380
Qy	1381	GGTCAAACTATTGCGCATTTGATCAAAAGGAAAGGAGCAAGAAAGCAATTTGAAACCA	1440
Db	1381	GGTCAAACTATTGCGCATTTGATCAAAAGGAAAGGAGCAAGAAAGCAATTTGAAACCA	1440
Qy	1441	GTCAATCATGACCAAAATGAGAACTTACAGTGGTCAAGTAAGCTTAAAGGCAAGAAAGCA	1500
Db	1441	GTCAATCATGACCAAAATGAGAACTTACAGTGGTCAAGTAAGCTTAAAGGCAAGAAAGCA	1500
Qy	1501	GAACACACTCAACCAATCAGAGACAGAAATGCTTTAACTTCACTAACACATACCTTGCCT	1560
Db	1501	GAACACACTCAACCAATCAGAGACAGAAATGCTTTAACTTCACTAACACATACCTTGCCT	1560
Qy	1561	GGTGACTTTGGTTGGTTGACAAAATGAAATGGCCCATGATTTGCTGGTTGGTAAAG	1620
Db	1561	GGTGACTTTGGTTGGTTGACAAAATGAAATGGCCCATGATTTGCTGGTTGGTAAAG	1620
Qy	1621	AATGGTTACCAATCTACCAATGCAAGTACTGCTGCTAAATGGGCAAGTCTCTGATTGG	1680
Db	1621	AATGGTTACCAATCTACCAATGCAAGTACTGCTGCTAAATGGGCAAGTCTCTGATTGG	1680
Qy	1681	TCAGAAACTGGCGAGCCAAAGTGGCACTCTCTATATAATTTACTAGTTTGGCAGCG	1740
Db	1681	TCAGAAACTGGCGAGCCAAAGTGGCACTCTCTATATAATTTACTAGTTTGGCAGCG	1740
Qy	1741	TCACCAATTCACACACCCGAAAGTAGCCCTCTCAGCAGAACTATGCACTTCCACTT	1800
Db	1741	TCACCAATTCACACACCCGAAAGTAGCCCTCTCAGCAGAACTATGCACTTCCACTT	1800
Qy	1801	GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACACCAATACTCTGTTGG	1860
Db	1801	GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACACCAATACTCTGTTGG	1860
Qy	1861	GGCACTGCAAGAACCCAGAACTCTGGGAAGCTGGTTCCAAAGCCTGCAAGATGGTCAA	1920
Db	1861	GGCACTGCAAGAACCCAGAACTCTGGGAAGCTGGTTCCAAAGCCTGCAAGATGGTCAA	1920
Qy	1921	CTGAGCCCACTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTCTTGGTGGGAAACG	1980
Db	1921	CTGAGCCCACTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTCTTGGTGGGAAACG	1980
Qy	1981	TTGAAAGAAAGACTTTCAGCGAGCGCTGAACTTTGGACTAA	2019
Db	1981	TTGAAAGAAAGACTTTCAGCGAGCGCTGAACTTTGGACTAA	2019

RESULT 2	AX137736	2019 bp	DNA	linear	PAT 30-MAY-2001
LOCUS	AX137736				
DEFINITION	Sequence 1 from Patent EP1077260.				
ACCESSION	AX137736				
VERSION	AX137736.1	GI:14273909			

KEYWORDS	Mice minute virus	Db	481	TGTAATGTGCAATACACACGCTGAAAGNATTAATACTAAGAGAAATAGCAGAGCAAT	540
SOURCE	Mice minute virus	Qy	541	GAGTGGTTACTCTACTTATTATAGCATAGCAAAACCAAAAGAGACTATACCAAGTGT	600
ORGANISM	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.	Db	541	GAGTGGTTACTCTACTTATTATAGCATAGCAAAACCAAAAGAGACTATACCAAGTGT	600
REFERENCE	1 Nueesch, J. and Rommelaere, J.	Qy	601	GTTCCTTTTGGAAACATGATGCTTACTATTTTAACTTAAAGAAATAAGCACTAGT	660
AUTHORS	Parvovirus ns1 variants	Db	601	GTTCCTTTTGGAAACATGATGCTTACTATTTTAACTTAAAGAAATAAGCACTAGT	660
TITLE	Patent: EP 1077260-A 1 21-FEB-2001;	Qy	661	CCACCAAGAGACGGAGGCTATTCTTAGCAGTGACTCTGGCTGGAAAACCTAACTTTTAA	720
JOURNAL	Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts (DE)	Db	661	CCACCAAGAGACGGAGGCTATTCTTAGCAGTGACTCTGGCTGGAAAACCTAACTTTTAA	720
FEATURES	Location/Qualifiers	Qy	721	AAAGAGGCGAGCGCCATCTAGTGACAACTATACACTGATGACATGCGGCCAGAAACG	780
source	1. .2019	Db	721	AAAGAGGCGAGCGCCATCTAGTGACAACTATACACTGATGACATGCGGCCAGAAACG	780
CDS	/organism="Mice minute virus"	Qy	781	GTGTAACACACAGTACCACTGGCAGGAACTTAAGCGGCGCAGAACTTCAAACTTAAAAA	840
	/mcl_type="unassigned DNA"	Db	781	GTGTAACACACAGTAACTGCGCAGGAACTTAAGCGGCGCAGAACTTCAAACTTAAAAA	840
	/db_xref="taxon:10794"	Qy	841	GAAAGTTCTATTAAAACTACACTTAAAGAGCTGGTGCATAAAAAGAGTAACTTCAACAGAG	900
	/note="Wildtype Parvovirus NS1"	Db	841	GAAAGTTCTATTAAAACTACACTTAAAGAGCTGGTGCATAAAAAGAGTAACTTCAACAGAG	900
	1. .2019	Qy	901	GACTGATGATGATGCGAGCAGACAGTTTACATTTGAAATGATGCTCAACAGAGTGGAGAA	960
	/note="unnamed protein product"	Db	901	GACTGATGATGATGCGAGCAGACAGTTTACATTTGAAATGATGCTCAACAGAGTGGAGAA	960
	/codon_start=1	Qy	961	AACTGCTGAAAAATACGCTAGAGATTTGTACACTTAACCTCTAGCCAGAACCAAAACAGCA	1020
	/protein_id="CAC39989.1"	Db	961	AACTGCTGAAAAATACGCTAGAGATTTGTACACTTAACCTCTAGCCAGAACCAAAACAGCA	1020
	/db_xref="GI:14273910"	Qy	1021	TTTGACTTAATTTAGAAAAAGCTGAAACACGACAACTAAACCACTTTTCACTGCCTGAC	1080
	/translation="MAGNAYSDEVLGATNWLKEKSNQEVFSFVFNENVLNKGDIQW	Db	1021	TTTGACTTAATTTAGAAAAAGCTGAAACACGACAACTAAACCACTTTTCACTGCCTGAC	1080
	NSYKLELQDELKSLQSGAETTWDOSEDMEWETTVDEMTKOVFIQFDSLKKCLPEVL	Qy	1081	ACAAGAGCCTGAGAAATTTTTCCTTTCATGGCTGGAACCTATGTTAAAGTTTCCCATGCT	1140
	NTKNIFFPGDNNFVQHEWKDQGWCHVLIGKDSQAOQKWRBQLANVYWSRLVTA	Db	1081	ACAAGAACCTGAGAAATTTTTCCTTTCATGGCTGGAACCTATGTTAAAGTTTCCCATGCT	1140
	CNVQLTPABRIKLREIAENENWLLTYKHQTKDYTKCVLFQNMIAFYFLTKKIS	Qy	1141	ATTGCTGTGTTTTTAAACAGACAGGAGCAAAAGAAATACTGTTTTTATTTGAGACCA	1200
	TSPPDGGYFLSSDQKTFLEKGERHVLVSLYDDMRPETVETVLTVAQETKGRRI	Db	1141	ATTGCTGTGTTTTTAAACAGACAGGAGCAAAAGAAATACTGTTTTTATTTGAGACCA	1200
	QTKQESIKITLKLWHKRVTSPEDDMMQDPSYIEMMAQPGENLLKNTLEICTLTL	Qy	1201	GCAGCAGAGGCAAAATCTATTATTGCAAGCCATAGCACAAGCAGTGGCAATGTTGTT	1260
	ARTKTAFLILEKAETSKLTNFSLPDTRCIRIPAHGWNVVKVCHAIICVLNRQGGKR	Db	1201	GCAGCAGAGGCAAAATCTATTATTGCAAGCCATAGCACAAGCAGTGGCAATGTTGTT	1260
	NTVLFHPASTGKSIIAQIAQVGNVCYNAANVFNPDCTNKNLIWVERAGNFGQ	Qy	1261	TGCTATAATGCGAGCAATGTAACCTTTTCACTTTTAACTGTTACCAAGAACTTGATT	1320
	QYNOKALCSGOTIRIDOKGSKOIEPTVPLMTNENITVVRIGCEEPHTOPIRD	Db	1261	TGCTATAATGCGAGCAATGTAACCTTTTCACTTTTAACTGTTACCAAGAACTTGATT	1320
	RMLNHLHTLGRDFGLVDKNWPMICAWLVKNGYQSTWASCAKNGKVPDSEWAE	Qy	1321	TGGGTAGAGAGAGCTGGTAACTTTTGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT	1380
	PKPTPINILGDSRFSFTPKSPISUNYALPLASDLEDLALPEWSTPTNTPVAGTAE	Db	1321	TGGGTAGAGAGAGCTGGTAACTTTTGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT	1380
	TQNTGEAGSKACQDQGLSPTWSEIEEDLRACFACPLKKDFSEPLND"	Qy	1381	GGTCAAACTATTTCGCATTCGATCAAAAGGAAAGGCGACAAACAGATTGAAACCAACCA	1440
ORIGIN		Db	1381	GGTCAAACTATTTCGCATTCGATCAAAAGGAAAGGCGACAAACAGATTGAAACCAACCA	1440
Query Match	99.9%; Score 2017.4; DB 6; Length 2019;	Qy	1441	GTTCATCATGACCAAAATGAGAACTTACAGTGGTTCAGAAATAGGCTGCGAAGAAAGACCA	1500
Best Local Similarity	100.0%; Pred. No. 0;	Db	1441	GTTCATCATGACCAAAATGAGAACTTACAGTGGTTCAGAAATAGGCTGCGAAGAAAGACCA	1500
Matches 2018; Conservative	0; Mismatches 1; Indels 0; Gaps 0;	Qy	1501	GAAACACACTCAACCAATCAGACAGAAATGCTTTAACTTCATCTAAACATACCTTGCCT	1560
1	ATGGCTGGAAATGCTTACTCTGATGAGTTTGGGAGCAACCACTGGTTAAAGGAAAAA	Db	1501	GAAACACACTCAACCAATCAGACAGAAATGCTTTAACTTCATCTAAACATACCTTGCCT	1560
1	ATGGCTGGAAATGCTTACTCTGATGAGTTTGGGAGCAACCACTGGTTAAAGGAAAAA	Qy	1561	GGTGTACTTTGGTTGGTTCACAAAATGAATGGCCCATGATTTGTGCTTGGTTGGTAAAG	1620
61	AGTAACAGGAGAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGAAAA	Db	1561	GGTGTACTTTGGTTGGTTCACAAAATGAATGGCCCATGATTTGTGCTTGGTTGGTAAAG	1620
61	AGTAACAGGAGAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGAAAA	Qy			
121	GATATCGGATGGAATAGTTTACAAAAAGAGCTGACGAGGAGCGAGCTGAAATCTTTACAA	Db			
121	GATATCGGATGGAATAGTTTACAAAAAGAGCTGACGAGGAGCGAGCTGAAATCTTTACAA	Qy			
181	CGAGGAGCGGAAACTACTTTGGACCAAAAGCGAGGACATGGAATGGGAAACCAAGTGGAT	Db			
181	CGAGGAGCGGAAACTACTTTGGACCAAAAGCGAGGACATGGAATGGGAAACCAAGTGGAT	Qy			
241	GAAATGACCAAAAGCAAGTATTCATTTTGTATTTCTTGGTTTAAAAAAATGTTTATTTGAA	Db			
241	GAAATGACCAAAAGCAAGTATTCATTTTGTATTTCTTGGTTTAAAAAAATGTTTATTTGAA	Qy			
301	GTGCTTAAACAAAGAAATATTTCTGCTGATGTTTAAATGGTTGTGCAACATGAATGG	Db			
301	GTGCTTAAACAAAGAAATATTTCTGCTGATGTTTAAATGGTTGTGCAACATGAATGG	Qy			
361	CGAAAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGACCTTTTACTCAAGCT	Db			
361	CGAAAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGACCTTTTACTCAAGCT	Qy			
421	CAAGGAAATGGTGGAGAGGCACTTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCC	Db			
421	CAAGGAAATGGTGGAGAGGCACTTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCC	Qy			
481	TGTAATGTGCAATACACACGCTGAAAGNATTAATACTAAGAGAAATAGCAGAGCAAT	Db			

QY	1621	AATGGTTTACCAATCTACCATGGCAAGTACTGTGCTAAATGGGCAAAAGTTCTCTGATTGG	1680
Db	1621	AATGGTTTACCAATCTACCATGGCAAGTACTGTGCTAAATGGGCAAAAGTTCTCTGATTGG	1680
QY	1681	TCAGAAAACCTGGCGGAGCAAAAGTGCCAACTCTATATAATTTACTAGTTCGGCAGCG	1740
Db	1681	TCAGAAAACCTGGCGGAGCAAAAGTGCCAACTCTATATAATTTACTAGTTCGGCAGCG	1740
QY	1741	TCACCAATTCAGCACCGAAAGTACGCTCTGAGCAGAACTATGCACTAATCTCACTT	1800
Db	1741	TCACCAATTCAGCACCGAAAGTACGCTCTGAGCAGAACTATGCACTAATCTCACTT	1800
QY	1801	GCATCGATCTCGAGGACCTGGCTTTAGAGCTTTGAGCACACCAATACTCTCTGTTGG	1860
Db	1801	GCATCGATCTCGAGGACCTGGCTTTAGAGCTTTGAGCACACCAATACTCTCTGTTGG	1860
QY	1861	GGCACTGCAGAAACCCAGAACACTCGGGAGAGCTGGTTCCAAAGCCTGCCAAGATGCTCAA	1920
Db	1861	GGCACTGCAGAAACCCAGAACACTCGGGAGAGCTGGTTCCAAAGCCTGCCAAGATGCTCAA	1920
QY	1921	CTGAGCCCACTGTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGGTGCAGAACCG	1980
Db	1921	CTGAGCCCACTGTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGGTGCAGAACCG	1980
QY	1981	TTGAAGAAAGACTTCAGCGAGCGCTGAACTTCGACTAA	2019
Db	1981	TTGAAGAAAGACTTCAGCGAGCGCTGAACTTCGACTAA	2019
RESULT 3			
LOCUS	PAMV2	5081 bp	DNA linear VRL 10-FEB-1999
DEFINITION	Minute virus of mice with two major open reading frames (genome).		
ACCESSION	V01115		
VERSION	V01115.1	GI:60911	
KEYWORDS	coat protein; genome; origin of replication; overlapping genes; terminal repeat.		
SOURCE	Minute minute virus		
ORGANISM	Mice minute virus		
REFERENCE	1 (bases 1 to 5081)		
AUTHORS	Astell,C.R., Thomson,M., Merchlinsky,M. and Ward,D.C.		
TITLE	The complete DNA sequence of minute virus of mice, an autonomous parvovirus		
JOURNAL	Nucleic Acids Res. 11 (4), 999-1018 (1983)		
MEDLINE	83143341		
PUBMED	6298737		
COMMENT	The messenger RNA of this virus (colinear with the strand listed) is spliced, but the exact splice sites are not known. The listed strand is complementary to the one which is included in the virion.		
FEATURES	Location/Qualifiers		
source	1..5081		
	/organism="Mice minute virus"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:10794"		
	114..2279		
	/notes="unlabeled protein product; coding sequence"		
	/codon_start=1		
	/protein_id="CAA24309.1"		
	/db_xref="GI:60912"		
	/db_xref="GOA:P03134"		
	/db_xref="Swiss-Prot:P03134"		
	/translation="MISGSGSLNKGAKRWKFWKVVYKOLLKSVTVLFFHSVSRDAQKE SNLTWAGNAYSDVLGATNMLKEKNOEVSFVKENVNOLNGKDIWNYSVKELQEDLSLQRAETWDSMEMETTVDTMKQVFPDLSLVKKCLFVLNKNIPFGD VNWVFOHWKQDQGHVLIIGKQFSAQGWKRWQLNYSWRLVATCNVQLTPAE RIKLREIAEDNEFWLLTYKHKQDYTKVLFGNMIAAYFLTKKKISTSPRGGY FLSSDGMKTNFLKEGERHLVSLYDDMRPETVETVTAQETKRGRIQTKKEYSIK TTLKELVHRVTSPELMMQPDYSLEMAQPGENLLKNTLEICTLTLARTKTAFLD ILEKETSCLTFLSPDTECTRIAPFHGNVYKVCALICVLNRGGKENTVLFHGA STGKSLIAQALQAVNGVCYNAANVFNPNCTNKNLIWSEAGNFGQVQVQFKAI C SQGTIRIDQKQKQIEPTPVTMTNENITVVRIGCEERPEHTQPIRDMRLNIHLTH		
CDS			
	/note="unlabeled protein product; coding sequence"		
	/codon_start=1		
	/protein_id="CAA24310.1"		
	/db_xref="GI:60913"		
	/db_xref="GOA:P03137"		
	/db_xref="Swiss-Prot:P03137"		
	/translation="MAPPAPRAKRGKLRDGLWLVY"		
	2332..2361		
	/note="unlabeled protein product; coding sequence"		
	/codon_start=1		
	/protein_id="CAA24311.1"		
	/db_xref="GI:584078"		
	/db_xref="TrEMBL:O89491"		
	/translation="MVGWMGINV"		
	2354..2398		
	/note="unlabeled protein product; coding sequence"		
	/codon_start=1		
	/protein_id="CAA24312.1"		
	/db_xref="GI:60914"		
	/db_xref="TrEMBL:O89818"		
	/translation="MFNYLFYRPEITWF"		
	2383..4551		
	/note="unlabeled protein product; Protein sequence is in conflict with the conceptual translation; coding sequence"		
	/codon_start=1		
	/protein_id="CAA24313.1"		
	/db_xref="GI:1335397"		
	/db_xref="GOA:O84207"		
	/db_xref="TrEMBL:O84207"		
	/translation="NHLVLGVVPPGCKYKLPNCNSLDQCEPTNPSDAAKEHDEAYDQY IKSQNPDIYFSAQDFIDQDKADMGKGVHYFFRTKRAFAPKLAIDSFPSTGV SADGRTTPPAYIFINQARAKKLTSSAAQSSQTSQDTSQDSGNVHSAARVERA ADPGCGGGGGGGGVSQYDQTHYRFLGQGWBITALATRLVHLNPKSEY CRIRVHTTDTSVKGNMAKDAHQIWTWPSLVDAWGVWLQPDQVQICNTMSQNL LVSLDQEIFNVVLKTVTDQILGQAIKIYNDLTACMVAVDNNILPVTPAANMET LGPYPKPTIASPYEYFCVDRDLSVYENQEGTVENVMGTPKGIPOFFETNTQOI TLIRTGDEFATGTYTFTNSVKLTITNQLRGQPPLLSTPEPDTAGTLTAGSR HGTTQMGVNWSEARITPAQVGFQCPHNDNFASRAGFAAPKPADITQGVDRKANG SVRYSYKQHGENSEWASHGPARYETWDSFSGSDTKDGFQSLAPLVVPPNLGIT NANPTGTNDIHFNSVNSYGLPTAFSPSPVYQCOIWKDELDEHKLPRHITAPFV CKNAPAGOMLVRLGPNLTDQYDNGATLSRIYVTGTFWKGKLTWRKLANTWNPV YQVSAEDNGNSYMSVTKLPLTATGNQVSLITRPFVARTY"		
ORIGIN			
Query Match	99.9%;	Score 2017.4;	DB 14; Length 5081;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 2018;	Conservative 0;	Mismatches 1;	Indels 0; Gaps 0;
QY	1	ATGGCTGGAATGCTTACTCTGATCAAGTTTGGGAGCAACCACTGGTTAAGGAAAA	60
Db	261	ATGGCTGGAATGCTTACTCTGATCAAGTTTGGGAGCAACCACTGGTTAAGGAAAA	320
QY	61	AGTAACCAAGAAAGTCTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAA	120
Db	321	AGTAACCAAGAAAGTCTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAA	380
QY	121	GATATCGGATGAATAGTTTACAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTTACAA	180
Db	381	GATATCGGATGAATAGTTTACAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTTACAA	440
QY	181	CGAGAGCGGAACTACTTGGACCAACCGAGGACATGGAATGGGAACACACAGTGGAT	240
Db	441	CGAGAGCGGAACTACTTGGACCAACCGAGGACATGGAATGGGAACACACAGTGGAT	500
QY	241	GAAATGACCAAAAAACAAGTATTCTATTTGATCTTTGGTTAAAAAATGTTTATTGAA	300
Db	501	GAAATGACCAAAAAACAAGTATTCTATTTGATCTTTGGTTAAAAAATGTTTATTGAA	560
QY	301	GTGCTTAACAAAGAAATATATTTCTCTGTGTAGTTAATTTGTTTGTGCAACATGATCG	360

Db 561 GTGCTTACACAAAGAAATATATTTCTGCTGATGTTAAATTTGGTTTGTGCAACATGAATGG 620
Qy 361 GGAAGAGACCAAGGCTGGCACTGCCATGTACTAAATTTGGAGGAAGGACATTTAGTCAAGCT 420
Db 621 GGAAGAGACCAAGGCTGGCACTGCCATGTACTAAATTTGGAGGAAGGACATTTAGTCAAGCT 680
Qy 421 CAAAGGAAATGGTGAGAAAGGCAACTAAATGTTTACTGGAGCAGATGGTGGTAAACAGCC 480
Db 681 CAAAGGAAATGGTGAGAAAGGCAACTAAATGTTTACTGGAGCAGATGGTGGTAAACAGCC 740
Qy 481 TGTAAATGCACTAACACACAGCTGAAGCAATTAACCTAAGAGAAATAGCAGAGCAAT 540
Db 741 TGTAAATGCACTAACACACAGCTGAAGCAATTAACCTAAGAGAAATAGCAGAGCAAT 800
Qy 541 GAGTGGGTTACTCTACTTACTTATAGCATAAGCAAAACCAAAAGACATATACCAAGTGT 600
Db 801 GAGTGGGTTACTCTACTTACTTATAGCATAAGCAAAACCAAAAGACATATACCAAGTGT 860
Qy 601 GTTCTTTTGGAAAACATGATGCTTACTTACTTATTTTAACTAAAGAAAATAGCACTAGT 660
Db 861 GTTCTTTTGGAAAACATGATGCTTACTTACTTATTTTAACTAAAGAAAATAGCACTAGT 920
Qy 661 CCACCAAGAGCGGAGGCTATTTCTTAGCAGTACTCTGGCTGGGAAACTAACTTTTAA 720
Db 921 CCACCAAGAGCGGAGGCTATTTCTTAGCAGTACTCTGGCTGGGAAACTAACTTTTAA 780
Qy 721 AAAGAGGCGGAGCGGCATCTAGTGAGCAAACTATACCTGATGACATGGGCGCAGAAAGC 740
Db 981 AAAGAGGCGGAGCGGCATCTAGTGAGCAAACTATACCTGATGACATGGGCGCAGAAAGC 1040
Qy 781 GTTGAACACCAAGTAAACCACTGCGCAGGAACTAAGCGCGGAGAAATCAAACTAAAAA 840
Db 1041 GTTGAACACCAAGTAAACCACTGCGCAGGAACTAAGCGCGGAGAAATCAAACTAAAAA 1100
Qy 841 GAACTTTCTATTAAGCACTTAAAGAGCTGGTGCATAAAGAGTAACTCAACAGAG 900
Db 1101 GAACTTTCTATTAAGCACTTAAAGAGCTGGTGCATAAAGAGTAACTCAACAGAG 1160
Qy 901 GACTGGATGATGATGAGCAGCAGACAGTTACATTTGAATGATGGCTCAACAGAGTGAGAA 960
Db 1161 GACTGGATGATGATGAGCAGCAGACAGTTACATTTGAATGATGGCTCAACAGAGTGAGAA 1220
Qy 961 AACCTGCTGAAAATAACGCTAGAGATTTGTACACTAACTCTAGCCAGAACCAAAACAGCA 1020
Db 1221 AACCTGCTGAAAATAACGCTAGAGATTTGTACACTAACTCTAGCCAGAACCAAAACAGCA 1280
Qy 1021 TTTGACTTAATTTAGAAAAGCTGAACAGCAAACTAACCAACTTTTCACTGCTGAC 1080
Db 1281 TTTGACTTAATTTAGAAAAGCTGAACAGCAAACTAACCAACTTTTCACTGCTGAC 1340
Qy 1081 ACAAGAGCTGCGAATTTTGTCTTTTCAATGCTGGAATATGTTTAAAGTTTGGCCATGCT 1140
Db 1341 ACAGAACCTGCGAATTTTGTCTTTTCAATGCTGGAATATGTTTAAAGTTTGGCCATGCT 1400
Qy 1141 ATTTGCTGTGTTTAAACAGACAGGAGCAAAAGAAATACGTGTTTATTTATGAGCA 1200
Db 1401 ATTTGCTGTGTTTAAACAGACAGGAGCAAAAGAAATACGTGTTTATTTATGAGCA 1460
Qy 1201 GCCAGACAGGCAATCTATTTATGCAAGCCATAGCAACAGCAGTTGGCAATGTTGT 1260
Db 1461 GCCAGACAGGCAATCTATTTATGCAAGCCATAGCAACAGCAGTTGGCAATGTTGT 1520
Qy 1261 TGTATAATGAGCCCAATGTAACCTTTTCAATTTAATGACTGTACCAACAGAACTTGAT 1320
Db 1521 TGTATAATGAGCCCAATGTAACCTTTTCAATTTAATGACTGTACCAACAGAACTTGAT 1580
Qy 1321 TGGGTAGAAAGAGCTGGTAACCTTTGACAGCAGTAAGCAAGCTTTAAAGCCATTTGCTCT 1380
Db 1581 TGGGTAGAAAGAGCTGGTAACCTTTGACAGCAGTAAGCAAGCTTTAAAGCCATTTGCTCT 1640
Qy 1381 GGTCAAACTATTGCGCATTTGATCAAAAGGAAAGGAGGAGCAAAACAGATTGAACCAACCA 1440

Db 1641 GGTCAAACTATTGCGCATTTGATCAAAAGGAAAGGAGGAGCAACAGATTGAACCAACCA 1700
Qy 1441 GTCATCATGACCAAAATGAGAACATTACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA 1500
Db 1701 GTCATCATGACCAAAATGAGAACATTACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA 1760
Qy 1501 GAAACACTCAACCAATCAGAGACAGAAATGCTTAACTTCACTACATACCTTTGGCCT 1560
Db 1761 GAAACACTCAACCAATCAGAGACAGAAATGCTTAACTTCACTACATACCTTTGGCCT 1820
Qy 1561 GGTGACTTTGGTTTGGTTGACAAAATGAATGGCCCATGATTTGTCTGGTTGGTAAAG 1620
Db 1821 GGTGACTTTGGTTTGGTTGACAAAATGAATGGCCCATGATTTGTCTGGTTGGTAAAG 1880
Qy 1621 AATGTTTACCAATCTACATGCAAGTACTGTCTTAAATGGGGCAAAAGTTCTCTGATTGG 1680
Db 1881 AATGTTTACCAATCTACATGCAAGTACTGTCTTAAATGGGGCAAAAGTTCTCTGATTGG 1940
Qy 1681 TCAGAAAATGGGGAGCCAAAGGTGCAACTCTCTATATAATTTACTAGCTTCGGCAGCG 1740
Db 1941 TCAGAAAATGGGGAGCCAAAGGTGCAACTCTCTATATAATTTACTAGCTTCGGCAGCG 2000
Qy 1741 TCACCAATTCACGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTTAACCTCCACTT 1800
Db 2001 TCACCAATTCACGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTTAACCTCCACTT 2060
Qy 1801 GCATCGGATCTCGAGGACCTGCTTTAGAGCTTTGGAGCAGACACCAAAATACTCTGTTGCG 1860
Db 2061 GCATCGGATCTCGAGGACCTGCTTTAGAGCTTTGGAGCAGACACCAAAATACTCTGTTGCG 2120
Qy 1861 GGCATGCGAAGAACCCAGAACACTGGGGAAGCTGTTCCAAAGCCTGCCAAGATGCTCAA 1920
Db 2121 GGCATGCGAAGAACCCAGAACACTGGGGAAGCTGTTCCAAAGCCTGCCAAGATGCTCAA 2180
Qy 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTCGGTGGGAAACCG 1980
Db 2181 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTCGGTGGGAAACCG 2240
Qy 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2019
Db 2241 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2279

RESULT 4

LOCUS 5149 bp ss-DNA linear VRL 22-MAY-1995
DEFINITION Minute virus of mice, complete genome.
ACCESSION J02275 M12521 M14704
VERSION J02275.1 GI:332293
KEYWORDS alternative splicing; capsid protein; complete genome; nonstructural protein.
SOURCE Mice minute virus
ORGANISM Mice minute virus
REFERENCE 1 (bases 1 to 5149)
AUTHORS Astell,C.R., Thomson,M., Merchlinsky,M. and Ward,D.C.
TITLE The complete DNA sequence of minute virus of mice, an autonomous parvovirus
JOURNAL Nucleic Acids Res. 11 (4), 999-1018 (1983)
MEDLINE 83143341
PUBMED 6298737
REFERENCE 2 (bases 1 to 5149)
AUTHORS Astell,C.R., Gardiner,E.M. and Tattersall,P.
TITLE DNA sequence of the lymphotropic variant of minute virus of mice, MVM(i), and comparison with the DNA sequence of the fibrotropic prototype strain
JOURNAL J. Virol. 57 (2), 656-669 (1986)
MEDLINE 86115415
PUBMED 3502703
REFERENCE 3 (sites)
AUTHORS Morgan,W.R. and Ward,D.C.
TITLE Three splicing patterns are used to excise the small intron common to all minute virus of mice RNAs

J. Virol. 60 (3), 1170-1174 (1986)
78061199
3783817
PUBMED
COMMENT
Original source text: Minute virus of mice (strain MVM(p)), passed in mouse 1 (variant A-9) cells.
The parvoviridae family contains two groups that infect mammalian hosts: (i) defective (helper-dependent) adeno-associated viruses, and (ii) autonomous (helper-independent) parvoviruses. MVM is a member of the latter group. Both groups have been demonstrated to package both plus and minus strands (in separate particles) of the ss-DNA genome, though the minus strand is more typically packaged in the latter group.
The sequence below corresponds to the plus (+) strand, also referred to as the C-strand. The minus (-) strand is also referred to as the V-strand.
The 3' and 5' termini both exhibit the potential for forming stable 'fold-back' hairpins; these sequences appear to play a role in replication [1].
The left and right halves of the genome encode two distinct, but overlapping transcriptional units. The transcripts can be summarized [1] (1 map unit (mu) = 51 bp):
R1 (4.8 kb): 4.5 mu - 46 mu; 46+ mu - 95 mu
R2 (3.3 kb): 4.5 mu - 10.7 mu; 38 mu - 46 mu; 46+ mu - 95 mu
R3 (3.0 kb): 40 mu - 46 mu; 46+ mu - 95 mu
R3 is the major transcript.
There are two major open reading frames, both on the plus (or C) strand. The left side ORF (261-2279) probably encodes a non-capsid protein of 85 kd; the right side ORF probably encodes the viral capsid proteins, VP1 (or A, 83 kd), VP2 (or B, 64 kd), and VP3 (or C, 61 kd). But because of uncertainties about the precise splice points in the transcripts, the exact starts, stops and (possible) intron boundaries are not known.
revision 4804 4870 a-65bp-a in [2]; aa in [1] [2]
revises [1].
[3] sites: splice sites.
Location/Qualifiers
1. .5149
/organism="Mice minute virus"
/mol_type="genomic DNA"
/strain="MVM(p)"
/db_xref="taxon:10794"
/lab_host="mouse 1 (variant A-9) cell"
114. .2279
/gene="NS1"
114. .2279
/gene="NS1"
/note="putative"
/codon_start=1
/product="nonstructural protein"
/protein_id="AAA67108.1"
/db_xref="GI:825481"
/translation="MISGSLNQAKRWKFKVYKOLLKSTVYLFHVSRSRDAOKE
SNQLTMAGNAYSDEVLGATNWLKESNQVFSFVKENNVOLNGKDIGNWSYKKELOE
DELKSLQRAETWQSEDMETVYDENTKKQVIFPSLVKKCLFEVLNKNIFPGD
VNWVQHEGQDQGHCHVLLIGGKQSDQAGKRWRLNYSRWLVLTACNVQLTPAE
RLKREIAEDNNEWHLTKYKQDQYKCVLFGNMIAYFLTKKISTSPRDGGY
FLSDGKWNFLKEGERHLVSKLYTDDMPETVETVTAQETKRGRIQTKKEVSIK
TTLKSLVHRVTSPEDMMQPSYIEMMAQPGENLLKNTLCTLTARTKTAFLD
ILEKETSCLTNFSLPDTTCRIFAFHGMVNVKVCALCCVLRQGGKNTVLFHGA
STGKSLIAQAQVNGVGCYNAANVFPNDCTNKLIIWEAEAGNFGQVNOFKAIC
SQGTRIDQKSGSKQIEPTVIMTNENITVVRIGCEERPEHTQPIRDMNLNHLTH
TLPGDFGLVDKNEWPMICAWLVNGYQSTMASYCAKWKVPDWSNWAEPKPTPINL
LGSARSPTTQKSLPLSQNYALTPLASDLLEDALEPWSPTNTPVAGTAETQNTGAGS
KACQDQQLSPTWSEIEEDLRACFGAELPKKDFSEPLNLD"
200. .2279
/gene="NS1"
261. .2279
/gene="NS1"
/codon_start=1
/product="nonstructural protein"
/protein_id="AAA67109.1"
/db_xref="GI:332295"
/translation="MAGNAYSDEVLGATNWLKESNQVFSFVKENNVOLNGKDIGNW

NSYKKELOEDELKSLQRAETWQSEDMETVYDENTKKQVIFPSLVKKCLFEVL
NTKNIFPGDWNVFWHEGKQDQGHCHVLLIGGKQSDQAGKRWRLNYSRWLVLTAC
NVQLTPAEIRIAEDNNEWHLTKYKQDQYKCVLFGNMIAYFLTKKISTSPRDGG
YFLSDGKWNFLKEGERHLVSKLYTDDMPETVETVTAQETKRGRIQTKKEVSIK
TTLKSLVHRVTSPEDMMQPSYIEMMAQPGENLLKNTLCTLTARTKTAFLD
ILEKETSCLTNFSLPDTTCRIFAFHGMVNVKVCALCCVLRQGGKNTVLFHGA
STGKSLIAQAQVNGVGCYNAANVFPNDCTNKLIIWEAEAGNFGQVNOFKAIC
SQGTRIDQKSGSKQIEPTVIMTNENITVVRIGCEERPEHTQPIRDMNLNHLTH
TLPGDFGLVDKNEWPMICAWLVNGYQSTMASYCAKWKVPDWSNWAEPKPTPINL
LGSARSPTTQKSLPLSQNYALTPLASDLLEDALEPWSPTNTPVAGTAETQNTGAGS
KACQDQQLSPTWSEIEEDLRACFGAELPKKDFSEPLNLD"
2002. .2398
/gene="vp"
2002. .2280
/gene="vp"
/note="major transcription start site"
/number=1
2006. .2280
/gene="vp"
/note="minor transcription start site"
/number=1
2009. .2280
/gene="vp"
/note="minor transcription start site"
/number=1
2281. .2398
/gene="vp"
/note="alternative intron"
2281. .2376
/gene="vp"
/note="alternative intron"
join(2286. .2316,2399. .4557)
/gene="vp1"
join(2286. .2316,2399. .4557)
/gene="vp1"
/codon_start=1
/protein_id="AAA67111.1"
/db_xref="GI:332294"
/translation="MAPPKRAKRGWPPGKYLGPGNSLDQGEPTNPSDAAKSHDE
AYDOYKSGKNPILYFSAADQRFIDQTKADKDWGKGVHYFRTKRAFAPKLTADSEP
GTSGVSRAGKTRPPAYIFINQARAKKLTSSAAQSSQTSWSDGTSQSDSGNAVHSA
KSENYCRIRVHNTDTSYKGNNAKDDAHEQIWTNPSLDANAGVWLPQSDQWYICNT
MSQLNLSDQIEFNVVLKTYEQDLGGQAIKIYNNDLTACMVAWVDSNNILPYTPAA
NSMETLGFYPMKPTIASPYVFCVDRDLSVTYENQEGTVEHNVMTGPKMNSQPFIT
ENTQOITLRTGDEPACTYFDTNSVKLTHTWOTNOLGOPPLLTSTPEADTDAGTL
TAQSRHGTQMGVNVWSEAIRTPAQVGFQCPNDPEASRAGPAAKVPADITQGV
DKEANGSVYISYKQHGHNWASHGAPAPERITWDETSPGSGRDTKDGFIQSAPLVVPPP
LNGILTNANPIGTKNIDHFSNVFNSYGLTAFSHSPSPYVPGQIWKDELDEHKPLRLH
ITAPFCVKNAPAGQMLVRLGNLTDQYDNGATLSRIVTYGTFFWKGLTWRAKLRLAN
TTWNPVYQVSAEDNGNSYMSVTKWLTATGNMQSVPLITRPFVARTY"
2286. .2354
/gene="vp"
/note="ORF1; putative"
/codon_start=1
/product="unknown protein"
/protein_id="AAA67110.1"
/db_xref="GI:825482"
/translation="MAPPKRAKRGKGLRDGLWLVGY"
2286. .2316
/gene="vp"
/number=1
2317. .2398
/gene="vp"
/note="vp intron (alt.)"
2332. .2361
/gene="vp"
/note="ORF3; putative"
/codon_start=1
/product="unknown protein"
/protein_id="AAA67112.1"
/db_xref="GI:825483"
/translation="MVGWVGINV"


```
Qy 1741 TCACCAATTCAGCACCGAAAGTACGCCTCTCAGCCAGAACATATGCACTAACTCCACTT 1800
Db 2001 TCACCAATTCAGCACCGAAAGTACGCCTCTCAGCCAGAACATATGCACTAACTCCACTT 2060
Qy 1801 GCATCGGATCTCAGGACCTGGCTTTAGAGCTTTGAGAGGACACCAAAATATCTCTGTTGGC 1860
Db 2061 GCATCGGATCTCAGGACCTGGCTTTAGAGCTTTGAGAGGACACCAAAATATCTCTGTTGGC 2120
Qy 1861 GGCACTCAGAAACCCAGAACCTGGGAGAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920
Db 2121 GGCACTCAGAAACCCAGAACCTGGGAGAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 2180
Qy 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGGCTGCTTCGGTGGGAACCG 1980
Db 2181 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGGCTGCTTCGGTGGGAACCG 2240
Qy 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGACTAA 2019
Db 2241 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGACTAA 2279

RESULT 5
AX137739
LOCUS AX137739 2019 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 4 from Patent EPI077260.
ACCESSION AX137739
VERSION AX137739.1 GI:14273913
KEYWORDS Mice minute virus
SOURCE Mice minute virus
ORGANISM Mice minute virus
REFERENCE 1
AUTHORS Nueesch,J. and Rommelaere,J.
TITLE Parvovirus ns1 variants.
JOURNAL Patent: EP 107260-A 4 21-FEB-2001;
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts
(DE)

FEATURES
source
1..2019
/organism="Mice minute virus"
/mol_type="unassigned DNA"
/db_xref="taxon:10794"
/notes="Parvovirus NS1 variant"
1..2019
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC39991.1"
/db_xref="GI:14273914"
/translation="MAGNAYSDVLGATNWLKESNOEVFSFVKENVQLNKDITGW
NSYKBLQDELKSLQGAETWDSQEDMEWETTVDEMTKQVFI FDSLKKCLFEVL
NTKNI PPGDVNWPVQHEWKGDOGWCHVLIGKDFSOAQGWRLQNLVYWSRLVTA
CNVQLTPAERIKLREIAEDNEWTLITYKHQTKDYTKCVLFQNMIAIYFLTKKIS
TSPROGGTFLSDSGKWNFLKEGERHLVSKLTDMDRPEVETVITTAQETKGRRI
QTKKEVAITLKLVLKRVTSPEMDMMQPDSEYEMMAOPGGENLLKNTLRICTLTL
ARTKFAFLILEKSETSLTNFSLPDRTRCIPAFHGMNVKVCNLI CCVLNRQGGKR
NTVLPFGPASTGKSI IAOIAOAVGNVGCNAAVNFNFNDCTNKNLI WVEBAGNFRQ
OVNQPKAI CSGOTRIRIDOKGSKSJOIEPTVIMTNTNENITVVRIGCEERPEHTQIRD
RLNIHLHTLPGDGLNDKNWPMICAMLVKNGVQSTWASCAKWKGVDPDSEWAE
PKVPTFNLGLSGARSFPTPKSTPLSQNTALPTLASDLDLALPEWSTNTPTVAGTAE
TQNTGAGSKACODGQLSPTWSEIEEDLRACFGAEPLKKDFSEPLND"

ORIGIN
Query Match 99.8%; Score 2015.8; DB 6; Length 2019;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGCTGGAATGCTTACTCTGATGAAGTITTTGGGAGCAACCACTGGTTAAAGGAAAA 60
Db 1 ATGGCTGGAATGCTTACTCTGATGAAGTITTTGGGAGCAACCACTGGTTAAAGGAAAA 60
Qy 61 AGTAACGAGGAAGTGTCTCATTTGTTTTTAAAAATGAAAAATGTTCAACTGAATGAAAA 120
```

```
Db 61 AGTAACGAGGAAGTGTCTCATTTGTTTTTAAAAATGAAAAATGTTCAACTGAATGAAAA 120
Qy 121 GATATCGGATGGAATAGTTTACAAAAAGAGAGCTGCAGAGGAGCAGAGCTGGAATCTTTACAA 180
Db 121 GATATCGGATGGAATAGTTTACAAAAAGAGAGCTGCAGAGGAGCAGAGCTGGAATCTTTACAA 180
Qy 181 CGAGGAGCGGAAACTACTTTGGGACCAAGCAGGACATGGAATGGGAAACCCACAGTGGAT 240
Db 181 CGAGGAGCGGAACTACTTTGGGACCAAGCAGGAGACATGGAATGGGAAACCCACAGTGGAT 240
Qy 241 GAAATGACCAAAAAAGCAAGTATTCATTTTGTGATCTTTTGGTTAAAAAATGTTTATTTGAA 300
Db 241 GAAATGACCAAAAAAGCAAGTATTCATTTTGTGATCTTTTGGTTAAAAAATGTTTATTTGAA 300
Qy 301 GTGCTTTAAACAAAAAGATATATTTCTGCTGATGTTAAATTTGTTGTCGCAACATGAATGG 360
Db 301 GTGCTTTAAACAAAAAGATATATTTCTGCTGATGTTAAATTTGTTGTCGCAACATGAATGG 360
Qy 361 GGAAAAAGACCAAGGCTGGCACTGCCATGTACTAAATTTGGAGGAAAGGACTTTTGTCAAGCT 420
Db 361 GGAAAAAGACCAAGGCTGGCACTGCCATGTACTAAATTTGGAGGAAAGGACTTTTGTCAAGCT 420
Qy 421 CAAGGAAATGTTGGAGAAAGGCAACTAAATGTTTACTGAGCAGATGTTGTTGTAACAGCC 480
Db 421 CAAGGAAATGTTGGAGAAAGGCAACTAAATGTTTACTGAGCAGATGTTGTTGTAACAGCC 480
Qy 481 TGTAAATGTGCAACTAACACCAGCTGAAAGAAATTAACAAAGAGAAATAGCAGAGCAAT 540
Db 481 TGTAAATGTGCAACTAACACCAGCTGAAAGAAATTAACAAAGAGAAATAGCAGAGCAAT 540
Qy 541 GAGTGGTGTACTCTACTTATTAAGCAATAAGCAAAACCAAAAGAGACTATATACCAGTGT 600
Db 541 GAGTGGTGTACTCTACTTATTAAGCAATAAGCAAAACCAAAAGAGACTATATACCAGTGT 600
Qy 601 GTTCTTTTGGAAAAATGATGTTGCTTACTATTTTAACTAAAAAGAAATAGACACTAGT 660
Db 601 GTTCTTTTGGAAAAATGATGTTGCTTACTATTTTAACTAAAAAGAAATAGACACTAGT 660
Qy 661 CCACCAAGAGACGAGGCTTATTTCTTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA 720
Db 661 CCACCAAGAGACGAGGCTTATTTCTTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA 720
Qy 721 AAAGAAGCGAGCGCCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCGAGAAACG 780
Db 721 AAAGAAGCGAGCGCCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCGAGAAACG 780
Qy 781 GTTGAACCAAGTAAACCACTGCGCAGGAACTAAGCGCGGAGAAATTCAAAATAAAAAA 840
Db 781 GTTGAACCAAGTAAACCACTGCGCAGGAACTAAGCGCGGAGAAATTCAAAATAAAAAA 840
Qy 841 GAAAGTTTCTATTAAACTACACTTAAAGAGCTGGTGCATATAAAGAGTAACTCACCAGAG 900
Db 841 GAAAGTTTCTATTAAACTACACTTAAAGAGCTGGTGCATATAAAGAGTAACTCACCAGAG 900
Qy 901 GACTGGATGATGATGACGACGACAGTGTACATTTGAAATGATGGCTCAACAGGTGGAGAA 960
Db 901 GACTGGATGATGATGACGACGACAGTGTACATTTGAAATGATGGCTCAACAGGTGGAGAA 960
Qy 961 AACCTGCTGAAAAATACCGTAGAGATTTGTACACTTAACCTAGCCAGAGAACCAAAACAGCA 1020
Db 961 AACCTGCTGAAAAATACCGTAGAGATTTGTACACTTAACCTAGCCAGAGAACCAAAACAGCA 1020
Qy 1021 TTTTGACTTAAATTTTAAAAAAGCTGAAACCAACCACTAACTTTTCACTGCTGAC 1080
Db 1021 TTTTGACTTAAATTTTAAAAAAGCTGAAACCAACCACTAACTTTTCACTGCTGAC 1080
Qy 1081 ACAAGAGCCTGCAAGATTTTGTCTTTTATGCTGGAACTATGTTTAAAGTTTGCCTAGCT 1140
Db 1081 ACAAGAGCCTGCAAGATTTTGTCTTTTATGCTGGAACTATGTTTAAAGTTTGCCTAGCT 1140
Qy 1141 ATTTGCTGTGTTTTTAAACAGACAGGAGGCAAAAGAAATACTGTTTATTTATCATGGACCA 1200
Db 1141 ATTTGCTGTGTTTTTAAACAGACAGGAGGCAAAAGAAATACTGTTTATTTATCATGGACCA 1200
```

```
QY 1201 GCAGACAGGCAAACTATTATTGACAAAGCCATAGCACAGCAGTTGGCAATGTTGGT 1260
DB 1201 GCAGACAGGCAAACTATTATTGACAAAGCCATAGCACAGCAGTTGGCAATGTTGGT 1260
QY 1261 TGCATAATGACGCAATGTAACCTTTTCCATTTTAATGACTGTACCAACAGAACTTGATT 1320
DB 1261 TGCATAATGACGCAATGTAACCTTTTCCATTTTAATGACTGTACCAACAGAACTTGATT 1320
QY 1321 TGGGTAGAAGAGCTGGTAACCTTTGGACAGCAAGTAAACAGATTTAAAGCAATTTGCTCT 1380
DB 1321 TGGGTAGAAGAGCTGGTAACCTTTGGACAGCAAGTAAACAGATTTAAAGCAATTTGCTCT 1380
QY 1381 GGTCAAACTATTGCGATTTGATCAAAAGAGAAAGGAGGAGCAAGATTTGAACCAACCA 1440
DB 1381 GGTCAAACTATTGCGATTTGATCAAAAGAGAAAGGAGGAGCAAGATTTGAACCAACCA 1440
QY 1441 GTCATCATGACCAAAATGAGACATTTACAGTGGTCAGATAGGCTGCCGAAGAAAGACCA 1500
DB 1441 GTCATCATGACCAAAATGAGACATTTACAGTGGTCAGATAGGCTGCCGAAGAAAGACCA 1500
QY 1501 GAACACACTCAACCAATCAGAGACAGAACTTTAAACATTCATCTAACACATACCTTGCCT 1560
DB 1501 GAACACACTCAACCAATCAGAGACAGAACTTTAAACATTCATCTAACACATACCTTGCCT 1560
QY 1561 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGCTTGGTTGGTAAAG 1620
DB 1561 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGCTTGGTTGGTAAAG 1620
QY 1621 AATGTTTACCAATCTACCATGCGCAAGCTACTGTGCTTAAATGGGGCAAGTTTCTGATTGG 1680
DB 1621 AATGTTTACCAATCTACCATGCGCAAGCTACTGTGCTTAAATGGGGCAAGTTTCTGATTGG 1680
QY 1681 TCAGAAAACTGGGCGGAGCCAAAGGTGCCAACTCTTATAAATTTACTAGTTTCGGCAGCG 1740
DB 1681 TCAGAAAACTGGGCGGAGCCAAAGGTGCCAACTCTTATAAATTTACTAGTTTCGGCAGCG 1740
QY 1741 TCACCAATTCAGCACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTAATCCACTT 1800
DB 1741 TCACCAATTCAGCACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTAATCCACTT 1800
QY 1801 GCATCGATCTCAGGACCTGGCTTTAGAGCTTTGAGAGCCTTGAGACACCAATATCTCTGTTGG 1860
DB 1801 GCATCGATCTCAGGACCTGGCTTTAGAGCCTTTGAGAGCCTTGAGACACCAATATCTCTGTTGG 1860
QY 1861 GGCACCTGCAGAAACCCAGACACTGGGGAAGCTGTTTCCAAAGCCTGCCAAGATGTCAT 1920
DB 1861 GGCACCTGCAGAAACCCAGACACTGGGGAAGCTGTTTCCAAAGCCTGCCAAGATGTCAT 1920
QY 1921 CTGAGCCCACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTGGTGGGAACCG 1980
DB 1921 CTGAGCCCACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTGGTGGGAACCG 1980
QY 1981 TTGAAGAAGACTTCAGGAGCGCTGAACTTGGACTAA 2019
DB 1981 TTGAAGAAGACTTCAGGAGCGCTGAACTTGGACTAA 2019
```

```
RESULT 6
AX137747
LOCUS AX137747 2019 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 12 from Patent EP1077260.
ACCESSION AX137747
VERSION AX137747.1 GI:14273921
KEYWORDS Mice minute virus
SOURCE Mice minute virus
ORGANISM Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE 1
AUTHORS Nueesch,J. and Rommelaere,J.
TITLE Parvovirus ns1 variants
JOURNAL Patent: EP 1077260-A 12 21-FEB-2001;
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts
```

(DE)

FEATURES
source

```
Location/Qualifiers
1..2019
/organism="Mice minute virus"
/mol_type="unassigned DNA"
/db_xref="taxon:10794"
/notes="Parvovirus.NS1 variant"
1..2019
/notes="unassigned protein product"
/codon_start=1
/protein_id="CAC39995.1"
/db_xref="GI:14273922"
/translation="MAGNAYSDEVLGATNWLKESNOEVFSFVKENVQLNGKDLGW
NYSKELQDELKSLQAGAEFTWDSQEDMEWETTKQVFIQDSLVKXKCLFEVL
NTKNLFFGDNVNFVQHEWQDQGHVLIQGGDFSAQKWKMRQUNLVYWSKLVIA
CNVQITPAERIKLREIAEDNWTLLTYKHQKDYTKCVLFVGNMIAFYFLTKKXIS
TSPRDDGGVFLSSDGWNTNFKLEGERHLVSKLYTDMRPETVETTTTAAQETKGR1
QTKKVS1KTLTKELVHKVTSPEWMMQPDYSY1EMMAQPGGKLNKLTLEICTLTL
ARTKTAFDLILEKAEKSLTNFSLPDRTRC1FAFHGWNVVKVCHAI1CVLNRQSGKR
NAVLFHPGASTGKSI1AQAQVNGVGCYNANVNFENDCTNKNLI1WEEAGNFGQ
QYNPKAICSGQTR1DQKQKSKQIEPTVPTVNTNENITVVRIGCEERPEHTQPIRD
RMLNTHLTHLPDFGLVDKNEWPMI1CAWLKNGYQSTWASVCAKWKVDPDMSNNAE
PKVPTPINLGSARGPFTPKPTPLSQNYALTPLASDLEDLALPEMSTPNTPVAGTAE
TQNTGEAGSKACQDGLSPTWSEIBEDLRACFGEPLKKDFSEPLND"
```

ORIGIN

```
Query Match 99.8%; Score 2015.8; DB 6; Length 2019;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCTGGAAATGCTTACTCTGATGAAAGTCTTGGGAGCAACCACTGTTTAAAGGAAAA 60
DB 1 ATGGCTGGAAATGCTTACTCTGATGAAAGTCTTGGGAGCAACCACTGTTTAAAGGAAAA 60
QY 61 AGTAACACGAGAGTGTCTCATTTGTTTAAAAATGAAATCTTCAACTGAATGAAAA 120
DB 61 AGTAACACGAGAGTGTCTCATTTGTTTAAAAATGAAATCTTCAACTGAATGAAAA 120
QY 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCGAGGAGCAAGCTGAAATCTTTACAA 180
DB 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCGAGGAGCAAGCTGAAATCTTTACAA 180
QY 181 CGAGGAGCGGAAACTACTTGGGACCAAGAGGAGCATGGAATGGGAAACCAACAGTGGAT 240
DB 181 CGAGGAGCGGAAACTACTTGGGACCAAGAGGAGCATGGAATGGGAAACCAACAGTGGAT 240
QY 241 GAAATGACCAAAAGCAAGTATTCTTTTGTATCTTGTGTTAAAAATCTTTTATTGAA 300
DB 241 GAAATGACCAAAAGCAAGTATTCTTTTGTATCTTGTGTTAAAAATCTTTTATTGAA 300
QY 301 GTGCTTAAACACAAAGAAATATATTTCTGCTGATGTTAAATTTGTTGCAACATGAATGG 360
DB 301 GTGCTTAAACACAAAGAAATATATTTCTGCTGATGTTAAATTTGTTGCAACATGAATGG 360
QY 361 GGAAGAGACCAAGCTGGCACTGCCATGTACTAATTTGGAGGAAAGCACTTTAGTCAAGCT 420
DB 361 GGAAGAGACCAAGCTGGCACTGCCATGTACTAATTTGGAGGAAAGCACTTTAGTCAAGCT 420
QY 421 CAAGGAAATGTTGGGAGAGGCACTAAATTTTACTGGAGCAGATGTTGTTAAACAGCC 480
DB 421 CAAGGAAATGTTGGGAGAGGCACTAAATTTTACTGGAGCAGATGTTGTTAAACAGCC 480
QY 481 TGTAAATGTCAACTAAACACAGCTGAAAGAAATTTAACTAAGAGAAATAGCAGAGACAAT 540
DB 481 TGTAAATGTCAACTAAACACAGCTGAAAGAAATTTAACTAAGAGAAATAGCAGAGACAAT 540
QY 541 GAGTGGTTTACTCTTACTTATAAGCATAGCAAAACCAAAAGACTATACCAAGTGT 600
DB 541 GAGTGGTTTACTCTTACTTATAAGCATAGCAAAACCAAAAGACTATACCAAGTGT 600
QY 601 GTTCTTTTGGAAACATGATTGCTTACTATTTTAACTAAAAAGAAATAAGCACTAGT 660
DB 601 GTTCTTTTGGAAACATGATTGCTTACTATTTTAACTAAAAAGAAATAAGCACTAGT 660
```

```
QY 661 CCACCAAGAGACGGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA 720
Db 661 CCACCAAGAGACGGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA 720
QY 721 AAAGAGGGGAGGCGCCATCTAGTGAGCAAACTATACACTGATGACATGGCGCCAGAAACG 780
Db 721 AAAGAGGGGAGGCGCCATCTAGTGAGCAAACTATACACTGATGACATGGCGCCAGAAACG 780
QY 781 GTTGAACACACAGTAACCTAGCGCAGGAACTAAGCGCGCAGAAATTCAAACTTAAAAA 840
Db 781 GTTGAACACACAGTAACCTAGCGCAGGAACTAAGCGCGCAGAAATTCAAACTTAAAAA 840
QY 841 GAAGTTTCTATTAAAACTACCTTAAAGAGCTGGTGCCATAAAAGAGTAACCTCACAGAG 900
Db 841 GAAGTTTCTATTAAAACTACCTTAAAGAGCTGGTGCCATAAAAGAGTAACCTCACAGAG 900
QY 901 GACTGATGATGATGACGAGCAGAGTACATTGAATGATGCTCAACAGAGTGAGAA 960
Db 901 GACTGATGATGATGACGAGCAGAGTACATTGAATGATGCTCAACAGAGTGAGAA 960
QY 961 AACCTGCTGAAAAATACGCTAGAGATTGTACACTAACTAGCCAGAACCAAAACAGCA 1020
Db 961 AACCTGCTGAAAAATACGCTAGAGATTGTACACTAACTAGCCAGAACCAAAACAGCA 1020
QY 1021 TTTGACTTAATTTTAGAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCTGAC 1080
Db 1021 TTTGACTTAATTTTAGAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCTGAC 1080
QY 1081 ACAGAGCCTGCGAAATTTTGGCTTTTTCATGGCTGGAATATGTTAAAGTTGCGCATGCT 1140
Db 1081 ACAGAGCCTGCGAAATTTTGGCTTTTTCATGGCTGGAATATGTTAAAGTTGCGCATGCT 1140
QY 1141 ATTGCTGTGTTTTAAACAGACAGAGCAAGAAAGAAATACCTGTTTTTATTCATGACCA 1200
Db 1141 ATTGCTGTGTTTTAAACAGACAGAGCAAGAAAGAAATACCTGTTTTTATTCATGACCA 1200
QY 1201 GCCAGACAGGCAAACTTATTATTCACAAGCCATAGCAAGCAGTGTGGCAATGTTGGT 1260
Db 1201 GCCAGACAGGCAAACTTATTATTCACAAGCCATAGCAAGCAGTGTGGCAATGTTGGT 1260
QY 1261 TGTATATAGCGAGCAATGTAACCTTTCAATTAATGACTGTACCAAGAACTTGATT 1320
Db 1261 TGTATATAGCGAGCAATGTAACCTTTCAATTAATGACTGTACCAAGAACTTGATT 1320
QY 1321 TGGTGAAGAGAGCTGTAACCTTTGACAGCAAGTAAACAGTTTAAAGCATTGCTCT 1380
Db 1321 TGGTGAAGAGAGCTGTAACCTTTGACAGCAAGTAAACAGTTTAAAGCATTGCTCT 1380
QY 1381 GGTCAAACTATTGCAATTGATCAAAAAGGAAAGGAGCAAAACAGATTGAACCAACCA 1440
Db 1381 GGTCAAACTATTGCAATTGATCAAAAAGGAAAGGAGCAAAACAGATTGAACCAACCA 1440
QY 1441 GTCATCATGACCAAAATGAGAACTTACAGTGGTGCAGATAGGCTGCGAAGAAAGACCA 1500
Db 1441 GTCATCATGACCAAAATGAGAACTTACAGTGGTGCAGATAGGCTGCGAAGAAAGACCA 1500
QY 1501 GAACACACTCAACCAATCAGACAGCAAGTCTTAACATTCATCAACATACATCTTGCT 1560
Db 1501 GAACACACTCAACCAATCAGACAGCAAGTCTTAACATTCATCAACATACATCTTGCT 1560
QY 1561 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTGCTTGGTTGGTGAAG 1620
Db 1561 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTGCTTGGTTGGTGAAG 1620
QY 1621 AATGGTTACCAATCTACATGCGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCTGATTGG 1680
Db 1621 AATGGTTACCAATCTACATGCGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCTGATTGG 1680
QY 1681 TCAGAAAACCTGGCGGAGCAAGGTGCCAACTCCTATATAATTTACTAGTTGCGGCACGC 1740
Db 1681 TCAGAAAACCTGGCGGAGCAAGGTGCCAACTCCTATATAATTTACTAGTTGCGGCACGC 1740
```

```
QY 1741 TCACCATTCAGACACCAGAAAGTAGCCTCTCAGCGACGAATATGCACTAATCCACTT 1800
Db 1741 TCACCATTCAGACACCAGAAAGTAGCCTCTCAGCGACGAATATGCACTAATCCACTT 1800
QY 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGCACACCAATATCTCTGTTGCG 1860
Db 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGCACACCAATATCTCTGTTGCG 1860
QY 1861 GGCACCTGCAGAAACCCAGAACACTGGGGAGCTGGTTCCAAAGCCTGCCAAGATGTCAA 1920
Db 1861 GGCACCTGCAGAAACCCAGAACACTGGGGAGCTGGTTCCAAAGCCTGCCAAGATGTCAA 1920
QY 1921 CTGAGCCCAACTTTGGTCAGAGATCGAGAGGATTTGAGAGCGTCTCGGTGCGGAACCG 1980
Db 1921 CTGAGCCCAACTTTGGTCAGAGATCGAGAGGATTTGAGAGCGTCTCGGTGCGGAACCG 1980
QY 1981 TTGAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019
Db 1981 TTGAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019

RESULT 7
AX137751
LOCUS AX137751 2019 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 16 from Patent EP1077260.
ACCESSION AX137751
VERSION AX137751.1 GI:14273925
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 Nueesch, J. and Rommelaere, J.
Parvovirus ns1 variants
Patent: EP 1077260-A 16 21-FEB-2001;
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts
(DE)
FEATURES
source
1..2019
/organism="Mice minute virus"
/mol_type="unassigned DNA"
/db_xref="taxon:10794"
/note="Parvovirus NS1 variant"
1..2019
/note="unnamed protein product"
/codon_start=1
/protein_id="C39997.1"
/db_xref="GI:14273926"
/translation="MAGNAYSDEVLGATNWLKEKSNQVSVFVFNENVLNGKDIQW
NSYKLEQDELSLQGAETTWDSQEDMEWETTVDMTKKQVIFDLSLVKKCLFEVL
NTKNTFPQDVNWFQHEWQDQGHVLIQKDFQSAQKWRRLQNTVYWRWLVT
CNVOLTPAERIKLRIADENWVLTLYKHOTKKDYTKVLFQNMIAVYFLTKKRI
TSPPRDGGVFLSSDSQWKTNFKGERHLVSKLITDDMPETVETVTTAOTKRGRI
QTKREVSITLTKELVHKRVTSPEDMMQPSYIEMMAQPGENLLKNTLEICTLTL
AKTAFDLILEKASTKLITNFSLPDTRCIFAFHGMNVKVCCHAICVCLNKGGR
NTVLFHGPASTKSIIAQIAQAVNGVCYNAANVPFNDCNKNKLIWVEEAGNGQ
QYNQPKATCSGOAIRIDQKGSQKEPTPVITMTNENITVVRIGCEERPEHTQPRD
RMLNTHLTLPDGLVDKNWPMICAWLVKNQYQSTWASCAKWKYKVPDNGENWAE
PKVPTPINLGSARSPFTPKSTPLSQNYALTPLASDLEDLALPEWSTNTPVAGTAE
TQNTGAGSKACQDQGLSPWSEIBEDLRACFGAEPLKKDFSEPLND"

Query Match 99.8%; Score 2015.8; DB 6; Length 2019;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAGGAAAA 60
Db 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAGGAAAA 60
QY 61 AGTACCCAGAGTGTCTCATTTGTTTAAAAATGAAAAATGTTCAACTGAATGAAAA 120
Db 61 AGTACCCAGAGTGTCTCATTTGTTTAAAAATGAAAAATGTTCAACTGAATGAAAA 120
```

QY	121	GATATCGGATGGAATAGTTACAAAAAGAGCTGCAGAGGACGAGCTGAAATCTTTACAA	180
DB	121	GATATCGGATGGAATAGTTACAAAAAGAGCTGCAGAGGACGAGCTGAAATCTTTACAA	180
QY	181	CGAGGAGCGGAAACTACTTCGGACCAAAGCGAGGACATGGAATGGGAAACCCACAGTGGAT	240
DB	181	CGAGGAGCGGAAACTACTTCGGACCAAAGCGAGGACATGGAATGGGAAACCCACAGTGGAT	240
QY	241	GAATGACCAAAAAAGCAAGTAATTCATTTTGTGTTGTTTAAAAAATGTTTATTTGAA	300
DB	241	GAATGACCAAAAAAGCAAGTAATTCATTTTGTGTTTAAAAAATGTTTATTTGAA	300
QY	301	GTGCTTAAACAAAGAAATATTTCTCTGGTGATGTTAAATTTGGTTTGTGCAACATGAATGG	360
DB	301	GTGCTTAAACAAAGAAATATTTCTCTGGTGATGTTAAATTTGGTTTGTGCAACATGAATGG	360
QY	361	GGAAAGACCAAGCTGGCACTGCCATGTACTAATTTGGAGGAAAGGACTTTTAGTCAAGCT	420
DB	361	GGAAAGACCAAGCTGGCACTGCCATGTACTAATTTGGAGGAAAGGACTTTTAGTCAAGCT	420
QY	421	CAAGGGAATTTGGTGGAGAGGCAACTAAATGTTTTACTTGGAGCAGATGGTTGGTAACAGCC	480
DB	421	CAAGGGAATTTGGTGGAGAGGCAACTAAATGTTTTACTTGGAGCAGATGGTTGGTAACAGCC	480
QY	481	TGTAATGTGAACCTAACACAGCTGAAAGAAATTTAACTAAGAGAAATAGCAGAAAGCAAT	540
DB	481	TGTAATGTGAACCTAACACAGCTGAAAGAAATTTAACTAAGAGAAATAGCAGAAAGCAAT	540
QY	541	GAGTGGTTTACTCTACTTACTTATAGCATAAGCAACCAAAAGAGCTATACCAAGTGT	600
DB	541	GAGTGGTTTACTCTACTTACTTATAGCATAAGCAACCAAAAGAGCTATACCAAGTGT	600
QY	601	GTTCTTTTGGAAACATGATGTTGCTTACTATTTTTTAACTAAAAAGAAAAATAAGCACTAGT	660
DB	601	GTTCTTTTGGAAACATGATGTTGCTTACTATTTTTTAACTAAAAAGAAAAATAAGCACTAGT	660
QY	661	CAACCAAGAGACGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA	720
DB	661	CAACCAAGAGACGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA	720
QY	721	AAAGAGGCGGAGCGCCATCTAGTGAGCAACTATACACTGATCAGTCGCGCCAGAAACG	780
DB	721	AAAGAGGCGGAGCGCCATCTAGTGAGCAACTATACACTGATCAGTCGCGCCAGAAACG	780
QY	781	GTTGAAACCAACAGTAACCTAGCGGAGGAACTAAGCGCGGAGAAATTCAAACTAAAAAA	840
DB	781	GTTGAAACCAACAGTAACCTAGCGGAGGAACTAAGCGCGGAGAAATTCAAACTAAAAAA	840
QY	841	GAAGTTTCTATTAAAACTACATTTAAAGAGCTGGTGCAAAAAAGAGTAACCTCACAGAG	900
DB	841	GAAGTTTCTATTAAAACTACATTTAAAGAGCTGGTGCAAAAAAGAGTAACCTCACAGAG	900
QY	901	GACTGGATGATGATGAGCAGCAGTATCATTTGAAATGATGCTCAACCCAGGTGAGAA	960
DB	901	GACTGGATGATGATGAGCAGCAGTATCATTTGAAATGATGCTCAACCCAGGTGAGAA	960
QY	961	AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTTAGCCAGAACCAAAACAGCA	1020
DB	961	AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTTAGCCAGAACCAAAACAGCA	1020
QY	1021	TTTGACTTAAATTTAGAAAAAGCTGAAACCAAGCAAACTAACCACTTTTCACTGCTGAC	1080
DB	1021	TTTGACTTAAATTTAGAAAAAGCTGAAACCAAGCAAACTAACCACTTTTCACTGCTGAC	1080
QY	1081	ACAGAGCTGCAAGATTTTGTCTTTTTCATGCTGGAATGATGTTAAAGTTTCCATGCT	1140
DB	1081	ACAAGAACCTGCAAGATTTTGTCTTTTTCATGCTGGAATGATGTTAAAGTTTCCATGCT	1140
QY	1141	ATTGCTGTGTTTTTAAACAGACAGGAGGCAAGAAATACTGTTTTTATTTTATGTGACCA	1200
DB	1141	ATTGCTGTGTTTTTAAACAGACAGGAGGCAAGAAATACTGTTTTTATTTTATGTGACCA	1200

QY	1201	GCAGCAGAGGCAAAATCTATTATTGCAAGCCATAGCACAAGCAGTTGGCAATGTTGGT	1260
DB	1201	GCAGCAGCAGGCAAAATCTATTATTGCAAGCCATAGCACAAGCAGTTGGCAATGTTGGT	1260
QY	1261	TGCTATAATCGACCAATGTAAACTTTTCCATTTTAACTGACTGTACCAAGAACTTGATT	1320
DB	1261	TGCTATAATCGACCAATGTAAACTTTTCCATTTTAACTGACTGTACCAAGAACTTGATT	1320
QY	1321	TGGGTAGAGAAAGCTGGTAACTTTTGGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT	1380
DB	1321	TGGGTAGAGAAAGCTGGTAACTTTTGGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT	1380
QY	1381	GGTCAAACTATTTCGCAATGATCAAAAGGAAAGGAGCAAAAGAGTTGAACCAACCA	1440
DB	1381	GGTCAAACTATTTCGCAATGATCAAAAGGAAAGGAGCAAAAGAGTTGAACCAACCA	1440
QY	1441	GTCAATGACCAAAATGAGAACATTTACAGTGGTCAAGATAGGCTCGGAGAAAGACCA	1500
DB	1441	GTCAATGACCAAAATGAGAACATTTACAGTGGTCAAGATAGGCTCGGAGAAAGACCA	1500
QY	1501	GAAACACATCAACCAATCGAGACAGAAATGCTTAAACATTCATCTAACACATACCTTGCCT	1560
DB	1501	GAAACACATCAACCAATCGAGACAGAAATGCTTAAACATTCATCTAACACATACCTTGCCT	1560
QY	1561	GGTGAATTTGGTTGGTTGACAAAAATGAATGSCCCATGATTTGTGCTTGGTTAAAG	1620
DB	1561	GGTGAATTTGGTTGGTTGACAAAAATGAATGSCCCATGATTTGTGCTTGGTTAAAG	1620
QY	1621	AATGGTTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGCAAAAGTTTCTGATTGG	1680
DB	1621	AATGGTTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGCAAAAGTTTCTGATTGG	1680
QY	1681	TCAGAAAACTGGCGGAGCCAAAGGTGCCAACTCTTATAAAATTTACTAGTTTGGCAGCG	1740
DB	1681	TCAGAAAACTGGCGGAGCCAAAGGTGCCAACTCTTATAAAATTTACTAGTTTGGCAGCG	1740
QY	1741	TCACCAATTCACGACACGAAAGATGACGCTCTCAGCAGAACTATGCACTAACCTCACTT	1800
DB	1741	TCACCAATTCACGACACGAAAGATGACGCTCTCAGCAGAACTATGCACTAACCTCACTT	1800
QY	1801	GCATCGATCTCGAGGACCTGGCTTTTAGAGCCTTTGGAGCACACCAATACTCTCTGTTGCG	1860
DB	1801	GCATCGATCTCGAGGACCTGGCTTTTAGAGCCTTTGGAGCACACCAATACTCTCTGTTGCG	1860
QY	1861	GGCACTGCAGAAACCCAGAACACTGGGGAAGCTGGTTTCCAAAGCTGCGCAAGATGGTCAA	1920
DB	1861	GGCACTGCAGAAACCCAGAACACTGGGGAAGCTGGTTTCCAAAGCTGCGCAAGATGGTCAA	1920
QY	1921	CTGAGCCCAACTTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGTTGGGAAACCG	1980
DB	1921	CTGAGCCCAACTTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGTTGGGAAACCG	1980
QY	1981	TTGAAGAAAGACTTTCAGCGAGCGCTGAACTGAGCTAA 2019	
DB	1981	TTGAAGAAAGACTTTCAGCGAGCGCTGAACTGAGCTAA 2019	

RESULT 8
MVMICG 5085 bp ss-DNA linear VRL 05-JUL-1995
LOCUS Minute virus of mice (MVM(i)), a lymphotropic variant of MVM,
DEFINITION complete genome.
ACCESSION M12032
VERSION M12032.1 GI:332289
KEYWORDS alternative splicing; capsid protein; complete genome;
nonstructural protein.
SOURCE Mice minute virus
ORGANISM Mice minute virus
REFERENCE 1 (bases 1 to 5085)
AUTHORS Åstell,C.R., Gardiner,E.M. and Tattersall,P.
TITLE DNA sequence of the lymphotropic variant of minute virus of mice,
MVM(i), and comparison with the DNA sequence of the fibrotropic

ORIGIN	1085 bp upstream of EcoRI site.	
Query Match	94.4%; Score 1905.4; DB 14; Length 5085;	
Best Local Similarity	96.5%; Pred. No. 0;	
Matches 1948; Conservative	0; Mismatches 71; Indels 0; Gaps 0;	
Qy	1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTTCGGAGCAACCACTGGTTTAAAGGAAAA 60	1021 TTTGACTTAATTTTAGAAAAAGCTGAACCAAGCAAACTAACCAACTTTTCTACTGCTGAC 1080
Db	262 ATGGCTGGAAATGCTTACTCTGATGAAGTTTTCGGAAACCAACCACTGGTTTAAAGGAAAA 321	1282 TTTGACTTGTATTTTAGAAAAAGCTGAACCAAGCAAACTAACCAACTTTTCTACTGCTGAC 1341
Qy	61 AGTAACCAAGGAAGTGTCTCATTTGTTTAAAAAATGAAATGTTCAACTGGAATGGAATA 120	1081 ACAAGAGCTGCGAAGATTTTGTCTTTTCATGGCTGGAACTATGTTTAAAGTTTGGCATGCT 1140
Db	322 AGTAACCAAGGAAGTGTCTCATTTGTTTAAAAAATGAAATGTTCAACTGGAATGGAATA 381	1342 ACAGAACTGCGAAGATTTTGTCTTTTCATGGCTGGAACTATGTTTAAAGTTTGGCATGCT 1401
Qy	121 GATATCGGATGGAATAGTTTACAAAAAGAGCTCGAGAGGAAGAGCTGGAATCTTTACAA 180	1141 ATTTGCTGTGTTTAAACACAGAGGAGGCAAAAGAAATACTGTTTATTTTCATGAGCA 1200
Db	382 GATATCGGATGGAATAGTTTACAAAAAGAGCTCGAGAGGAAGAGCTGGAATCTTTACAA 441	1402 ATTTGCTGTGTTTAAACACAGAGGAGGCAAAAGAAATACTGTTTATTTTCATGAGCA 1461
Qy	181 CGAGGAGCGAATCTTGGGACCAAGCGAGGACATGGAATGGGAAACCAACAGTGGAT 240	1201 GCAGACACAGGCAAAATCTATTTATGCAAGCCATAGCAAGAGCTTGGCAATGTTGCT 1260
Db	442 CGAGGAGCGAATCTTGGGACCAAGCGAGGACATGGAATGGGAAATCTACAGTGGAT 501	1462 GCAGTACAGGCAAAATCTATTTATGCAAGCCATAGCAAGAGCTTGGTAAATGTTGCT 1521
Qy	241 GAAATGACCAAAAAAGCAAGTATTCATTTTGTGTTTAAATAATGTTTATTTTGA 300	1261 TGTATAATGCGAGCAATGTAAACTTTTCCATTTTAAATGACTGTACCAACAGAACTTGAT 1320
Db	502 GAAATGACCAAAAAAGCAAGTATTCATTTTGTGTTTAAATAATGTTTATTTTGA 561	1522 TGTATAATGCGAGTAAATGTGAATTTTCCATTTTAAATGACTGTACCAACAGAACTTGAT 1581
Qy	301 GTGCTTTAACCAAAAGAAATATTTCTCTGGTGTATTTAAATTTGTTTGTGCAACATGAATG 360	1321 TGGGTAGAAAGAGCTGCTAACTTTTGGACAGCAAGTAAGCCAGTTTAAAGCCATTTGCTCT 1380
Db	562 GTGCTTTAGCAAAAAATATAGCTCTGCTGTATTTACTTTGTTTGTGAGCATGAATG 621	1582 TGGGTAGAAAGAGCTGCTAACTTTTGGACAGCAAGTAAGCCAGTTTAAAGCCATTTGCTCT 1641
Qy	361 GGAAGAGCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAGAGCTTTAGTCAAGCT 420	1381 GGTCAAACTATTTCGCATTGATCAAAAAGGAAAGGAGGAGCAAAAAGATTTGAACCAACCA 1440
Db	622 GGAAGAGCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAGAGCTTTAGTCAAGCT 681	1642 GGTCAAACTATTTCGCATTGATCAAAAAGGAAAGGAGGAGCAAAAAGATTTGAACCAACCA 1701
Qy	421 CAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGTTGGTAAACAGCC 480	1441 GTCATCATGACCAAAATGAGACATTTAGAGCAAGTAAGCCAGTTTAAAGCCATTTGCTCT 1500
Db	682 CAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGTTGGTAAACAGCC 741	1702 GTCATCATGACCAAAATGAGACATTTAGAGCAAGTAAGCCAGTTTAAAGCCATTTGCTCT 1761
Qy	481 TGTAAATGTGCACATAACACAGCTGAAAGAAATTTAACTAAGAGAAATAGCAGAGCAAT 540	1501 GAAACACATCAACCAATCAGAGACAGAAATGCTTAAACATTTCTAAACATACATCTTGCCT 1560
Db	742 TGTAAATGTGCACATAACACAGCTGAAAGAAATTTAACTAAGAGAAATAGCAGAGCAAT 801	1762 GAAACACATCAACCAATCAGAGACAGAAATGCTTAAACATTTCTAAACATACATCTTGCCT 1821
Qy	541 GAGTGGTTTACTTACTTATAGCATAAGCAAAACCAAAAGAGCTATACCAAGTGT 600	1561 GGTGACTTTGTTTGGTTGTTGACAAAAATGAATGGCCCATGATTTTGGTTGGTTAAAG 1620
Db	802 GAGTGGTTTACTTACTTATAGCATAAGCAAAACCAAAAGAGCTATACCAAGTGT 861	1822 GGTGACTTTGTTTGGTTGTTGACAAAAATGAATGGCCCATGATTTTGGTTGGTTAAAG 1881
Qy	601 GTTCTTTTGGAACTAGTGTCTTACTTATTTTAACTAAGAGAAATTAAGCACTAGT 660	1621 AATGGTTACCAATCTACCATGCAAGCTACTGTCTTAAATGGGGCAAAAGTTCTGTATTGG 1680
Db	862 GTTCTTTTGGAACTAGTGTCTTACTTATTTTAACTAAGAGAAATTAAGCACTAGT 921	1882 AATGGTTACCAATCTACCATGCAAGCTACTGTCTTAAATGGGGCAAAAGTTCTGTATTGG 1941
Qy	661 CCACCAAGAGAGCGAGGCTATTTTCTTAGCAGTGAATCTGGCTGGGAAACCTAATCTTTTA 720	1681 TCAGAAAACTGGGCGGAGCAAAAGTGCCTTCTTAAATTTTACTAGGTTTCGGCAGCG 1740
Db	922 CCACCAAGAGAGCGAGGCTATTTTCTTAGCAGTGAATCTGGCTGGGAAACCTAATCTTTTA 981	1942 TCAGAAAACTGGGCGGAGCAAAAGTGCCTTCTTAAATTTTACTAGGTTTCGGCAGCG 2001
Qy	721 AAGAGAGGAGCGGCACTAGTGAGCAACTATATACATGATGACATGCGGCCAGAAACG 780	1741 TCACCAATTACGACACCGAAAAAGTACGCTCTTCAGCCAGAACTATGCACTTAACTTCACTTT 1800
Db	982 AAGAGAGGAGCGGCACTAGTGAGCAACTATATACATGATGACATGCGGCCAGAAACG 1041	2002 TCACCAATTACGACACCGAAAAAGTACGCTCTTCAGCCAGAACTATGCACTTAACTTCACTTT 2061
Qy	781 GTTGAACCACTAGTACCACTGCGGAGGAACCTAAGCGGCGAGATTCAACTTAAATA 840	1801 GCATCGGATCTCGAGAGACCTGGCTTTAGAGCCTTGGAGCAACCAATATCTCTCTGTTGG 1860
Db	1042 GTTGAACCACTAGTACCACTGCGGAGGAACCTAAGCGGCGAGATTCAACTTAAATA 1101	2062 GCATCGGATCTCGAGAGACCTGGCTTTAGAGCCTTGGAGCAACCAATATCTCTCTGTTGG 2121
Qy	841 GAAATTTCTTAAAACTACATTTAAAGAGCTGGTGATAAAAAGAGTAACTCACCAGAG 900	1861 GGCATCTGAGAAACCCAGAAACCTGCGGGAAGCTGGTTTCAAAAGCTTGCCTTAAAGTGGTCAA 1920
Db	1102 GAGTTTCTTAAAAACCACTTAAAGAGCTAGTGCATTAAGAGATTAACCTCACCAGAA 1161	2122 GGCATCTGAGAAACCCAGAAACCTGCGGGAAGCTGGTTTCAAAAGCTTGCCTTAAAGTGGTCAA 2181
Qy	901 GACTGGATGATGTCAGGCAGACAGTTACATTTGAATGATGCTCAACCGAGTGGAGAA 960	1921 CTGAGCCCCAATTTGGTTGAGAGATTCGAGGAGGATTTGAGAGCGTGTCTTCGTTGCGGAACCG 1980
Db	1162 GACTGGATGATGTCAGGCAGACAGTTACATTTGAATGATGCTCAACCGAGTGGAGAA 1221	2182 CTGAGCCCCAATTTGGTTGAGAGATTCGAGGAGGATTTGAGAGCGTGTCTTCGTTGCGGAACCG 2241
Qy	961 AACCTGCTGAAAAATAGCTTAGAGATTTGTACATACTAGCTAGCAGAACCAAAACAGCA 1020	1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2019
Db	1222 AACCTGCTGAAAAATAGCTTAGAGATTTGTACATACTAGCTAGCAGAACCAAAACAGCA 1281	2242 TTGAAGAGAGACTTCAGCGAGCGCTGAACTTGGACTAA 2280

RESULT 9
MOU34253
LOCUS

MOU34253

4764 bp

DNA

linear

VRL 21-AUG-1996

DEFINITION	Mouse parvovirus 1b DNA.
ACCESSION	U34253
VERSION	U34253.1 GI:1464793
KEYWORDS	
SOURCE	Mouse parvovirus 1b
ORGANISM	Mouse parvovirus 1b
REFERENCE	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
AUTHORS	1 (bases 1 to 4764) Besselsen, D.G., Pintel, D.J., Purdy, G.A., Besch-Williford, C.L., Franklin, C.L., Hook, R.R., Jr. and Riley, L.K.
TITLE	Molecular characterization of newly recognized rodent parvoviruses
JOURNAL	J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
MEDLINE	96201434
PUBMED	8609486
REFERENCE	2 (bases 1 to 4764) Besselsen, D.G.
AUTHORS	Direct Submission
TITLE	Submitted (17-AUG-1995) David G. Besselsen, Department of
JOURNAL	Veterinary Pathology, University of Missouri-Columbia, W213
FEATURES	Veterinary Medicine Building, Columbia, MO 65211, USA Location/Qualifiers 1..4764 /organism="Mouse parvovirus 1b" /mol_type="genomic DNA" /specific_host="Mus musculus" /db_xref="taxon:42841"
ORIGIN	
Query Match	94.2%; Score 1902.2; DB 14; Length 4764;
Best Local Similarity	96.4%; Pred. No. 0;
Matches 1946; Conservative	0; Mismatches 73; Indels 0; Gaps 0;
QY	1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 60
Db	122 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 181
QY	61 AGTAACGAGGAAGTGTCTCATTTGTTTTTAAAAATGTAATGTTCAACTGAATGAAAA 120
Db	182 AGTAACGAGGAAGTGTCTCATTTGTTTTTAAAACTGAGATGTTCAACTGAATGAAAA 241
QY	121 GATATCGGATGGATAGTTTACAAAAAGAGCTGCAGGAGCAGCTGAAATCTTTACAA 180
Db	242 GATATCGGATGGAAATTAATACAAAAAGGAGCTGCAGGAGCAGCTGAAATCTTTACAA 301
QY	181 CGAGGAGCGGAATCTACTTGGGACCAAGCGAGGACATGGAATGGGAAACCACTGGAT 240
Db	302 CGAGGAGCGGAATCTACTTGGGACCAAGCGAGGACATGGAATGGGAACTACAGTGA 361
QY	241 GAAATGACCAAAAGCAAGTATTCATTTGTTGTTTGTGTTAAAAAATGTTTATTTGAA 300
Db	362 GAAATGACCAAAAGCAAGTATTCATTTGTTGTTTGTGTTAAAAAATGTTTATTTGAA 421
QY	301 GTGCTTAACACAAAGATATATTTCTGCTGATGTTAAATGTTTGTGCAACATGATGG 360
Db	422 GTGCTTAGCACAAAGATATATGCTCTGCTGATGTTTACTTGTGTTGTGAGCATGAATG 481
QY	361 GGAAGAGACCAAGGCTGGCACTGCCATGTACTTAATGGAGGAAGACATTTAGTCAAGCT 420
Db	482 GGAAGAGACCAAGGCTGGCACTGCCATGTACTTAATGGAGGAAGACATTTAGTCAAGCT 541
QY	421 CAAAGGAAATGTTGGAGAGGCAATTAATGTTTACTGGAGCAGATGGTTGTTAAACAGCC 480
Db	542 CAAAGGAAATGTTGGAGAGGCAATTAATGTTTACTGGAGCAGATGGTTGTTAAACAGCC 601
QY	481 TGTATGTGCAACTACACAGCTGAAAGATTAATTAAGTAAGTAAGTAAGTAAGTAAGTA 540
Db	602 TGTATGTGCAACTACACAGCTGAAAGATTAATTAAGTAAGTAAGTAAGTAAGTAAGTA 661
QY	541 GAGTGGGTTACTTCTACTTACTTATAAGCATTAAGCAACCAAAAAAGACTATATACCAAGTGT 600
Db	662 GAGTGGGTTACTTCTACTTACTTATAAGCATTAAGCAACCAAAAAAGACTATATACCAAGTGT 721
QY	601 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTTAAATAAGAAATAAGCACTAGT 660
Db	722 GTTCTTTTGGAAACATGATGCTTACTATTTTAAACCAAGAAATAAGCACTAGT 781
QY	661 CCACCAAGAGACGGAGGCTATTTCTTAGCAGTACTCTGGCTGGAAACTACTTTTAA 720
Db	782 CGCCCAAGGAGACGGAGGCTATTTCTTAGCAGTACTCTGGCTGGAAACTACTTTTAA 841
QY	721 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATCGGCGCAGAAACG 780
Db	842 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATCGGCGCAGAAACG 901
QY	781 GTTGAAACCAACAGTAACTACCTGCGCAGGAAACTTAAGCGCGGAGAAATCAAACTAAAAA 840
Db	902 GTTGAAACCAACAGTAACTACCTGCGCAGGAAACTTAAGCGCGGAGAAATCAAACTAAAAA 961
QY	841 GAAGTTTCTATTAAACTACACTTAAGAGCTGTTGTCATTAAGAGTAACCTCACCAGAG 900
Db	962 GAGGTTTCTATTAAACCACTTAAGAGCTGTTGTCATTAAGAGTAACCTCACCAGAA 1021
QY	901 GACTGATGATGATGAGCAGCAGACAGTTCATTTGAAATGATGCTCAACAGCTGGAGAA 960
Db	1022 GACTGATGATGATGAGCAGCAGACAGTTCATTTGAAATGATGCTCAACAGCTGGAGAA 1081
QY	961 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTTAACCTCTAGCCAGAAACCAACAGCA 1020
Db	1082 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTTAACCTCTAGCTAGAAACCAACAGCA 1141
QY	1021 TTTGACTTAAATTTAGAAAAAGCTGAAACCAACCACTTAACCACTTTTCACTGCCTGAC 1080
Db	1142 TTTGACTTAAATTTAGAAAAAGCTGAAACCAACCACTTAACCACTTTTCACTGCCTGAC 1201
QY	1081 ACAAGAGCTGCAGAAATTTTCTGCTTTTCATGCTGGAACTATGTTTAAAGTTTGCCATGCT 1140
Db	1202 ACAAGAGCTGCAGAAATTTTCTGCTTTTCATGCTGGAACTATGTTTAAAGTTTGCCATGCT 1261
QY	1141 ATTGCTGTGTTTTTAAACAGACAGAGGCAAAAGAAATACCTGTTTATTTATGAGACCA 1200
Db	1262 ATTGCTGTGTTTTTAAACAGACAGAGGCAAAAGAAATACCTGTTTATTTATGAGACCA 1321
QY	1201 GCAGCAGACAGGCAAAATCTATTATTCACAAGCCATAGCACAAGCTTGGCAATGTTGTT 1260
Db	1322 GCAGCAGACAGGCAAAATCTATTATTCACAAGCCATAGCACAAGCTTGGCAATGTTGTT 1381
QY	1261 TGCTATAATGACGCAATGTAAACTTTCCATTTAATGACTGTACCAACAGAACTTGATTT 1320
Db	1382 TGCTATAATGACGCAATGTAAACTTTCCATTTAATGACTGTACCAACAGAACTTGATTT 1441
QY	1321 TGGGTAGAAGAGCTGGTTAACTTTGGAACAGCAAGTAAACAGCTTTTAAAGCCATTTGCTCT 1380
Db	1442 TGGGTAGAAGAGCTGGTTAACTTTGGAACAGCAAGTAAACAGCTTTTAAAGCCATTTGCTCT 1501
QY	1381 GGTCAAACTATTTCGCATTTGATCAAAAGGAAAGGAGCAGCAAAACAGATTGCAACCAACCA 1440
Db	1502 GGTCAAACTATTTCGCATTTGATCAAAAGGAAAGGAGCAGCAAAACAGATTGCAACCAACCA 1561
QY	1441 GTCATCATGACCAAAAATGAGAACTTACAGTGGTCAAGTAAAGCTTGGCGGAGAAAGACCA 1500
Db	1562 GTCATCATGACCAAAAATGAGAACTTACAGTGGTCAAGTAAAGCTTGGCGGAGAGACCA 1621
QY	1501 GAAACACTCAACCAATCAGACAGAGATGCTTAACTTCACTTAACATACATACCTTTGCCT 1560
Db	1622 GAAACACTCAACCAATCAGACAGAGATGCTTAACTTCACTTAACATACATACCTTTGCCT 1681
QY	1561 GGTGACTTTGTTTGGTTGACAAAAATGAATGGCCCATGATTTGCTTTGTTGGTAAAG 1620
Db	1682 GGTGACTTTGTTTGGTTGACAAAAATGAATGGCCCATGATTTGCTTTGTTGGTAAAG 1741
QY	1621 AATGTTTACCAATCTACCAATGCAAGCTACTGTGCTTAAATGGGGCAAAAGTTCTCTGATGG 1680
Db	1742 AATGTTTACCAATCTACCAATGCAAGCTACTGTGCTTAAATGGGGCAAAAGTTCTCTGATGG 1801
QY	1681 TCAGAAAACTGGGCGGAGCAAGGTGCAACTCTTAAATTTTAACTTAAATTTAGTTTGGCAGCC 1740

Db 1802 ACAGAAAACCTGGCGGAGCCAGAGTCCGACTCCTATATAAATCACTAGTTCGGCACGC 1861

Qy 1741 TCACCAATTCACGACACCGAAAAGTAGCCTCTCAGCAGAACATATGCACTAACTCCACTT 1800

Db 1862 TCACCAATTCACGACACCGAAAAGTAGCCTCTCAGCAGAACATATGCACTAACTCCACTT 1921

Qy 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACACCAATACTCCTGTTGG 1860

Db 1922 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACACCAATACTCCTGTTGG 1981

Qy 1861 GGCACCTGCAGAAACCCAGAACACTGGGAAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920

Db 1982 GGCACCTGCAGAAACCCAGAACACTGGGAAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 2041

Qy 1921 CTGAGCCCAACTGGTTCAGAGATCGAGAGAGATTTAGAGCGTCTTCGTCGCGAACC 1980

Db 2042 CTGAGCCCAACTGGTTCAGAGATCGAGAGAGATTTAGAGCGTCTTCGTCGCGAACC 2101

Qy 1981 TTGAAGAGACTTCAGGAGCGCTGAACCTGCACTAA 2019

Db 2102 TTGAAGAGACTTCAGGAGCGCTGAACCTGCACTAA 2140

RESULT 10

PAMVMI
LOCUS
DEFINITION Mouse parvovirus minute virus immunosuppressive variant genome (= MvM1).
ACCESSION X02481
VERSION X02481.1 GI:60918
KEYWORDS coat protein; genome; origin of replication; overlapping genes; terminal repeat; unidentified reading frame.
SOURCE Mouse minute virus
ORGANISM Mouse minute virus
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE 1 (bases 1 to 5087)
AUTHORS Sahli,R., McMaster,G.K. and Hirt,B.
TITLE DNA sequence comparison between two tissue-specific variants of the autonomous parvovirus, minute virus of mice
JOURNAL Nucleic Acids Res. 13 (10), 3617-3633 (1985)
MEDLINE 85242059
PUBMED 3855242
COMMENT The genomes of MvMp and MvMi (immunosuppressive variant) have more than 96% of their sequence in common.
Data kindly reviewed (18-JUL-1986) by G. McMaster.

FEATURES

source 1..5087
/organism="Mouse minute virus"
/mol_type="genomic DNA"
/db_xref="taxon:10794"
misc_feature 1..116
/note="terminal hairpin region"
promoter 176..183
/note="TATA box"
gene 2405..4558
/gene="vp1"
CDS <2405..4558
/gene="vp1"
/function="coat protein"
/codon_start=1
/product="vp1 protein"
/protein_id="CAB46507.1"
/db_xref="GI:5419928"
/db_xref="GOA:P07302"
/db_xref="Swiss-Prot:P07302"
/translation="VPPGYLGGNSLDQGEPTNPSDAAKEHDEAYDQYIKSGKNP
YLVPSAADORFIDOTKADKWGGKGVHYFFRTKRAFAPKLATSEPTSGVSRAGKRT
RPPAYIFINQARAKKLTSSAAQCSOTMSDGTSDQPDGNGVHSAARVERAADPGGS
GGGSGGGGVSTGSDNQTHYRFLPDGHWVEITATATRLVHLNMPKSENYCRIRVHN
TTDTSVKNMAKODAHQEIWTPSLVDANAGVWLQPSDQWYICNTMSQLNLSLDOE
IFNVLKTVEQSDGQAIKIYNNDLTACMVAVDSNNILPYPYPAANSMTETLGFYPMK
PTIASPYRYFCVDRDLSTVYENQEGTIEHNVMGTPKGMNSQFFTIENQOITLLRTG

DEFACTGYVFDNTPVKLTHTWOTNRQLGQPLLSTFPPEADTDAGTLTAQGSRRHATQM
EYVWSEAIRTPAOCVPCOPNDPEASAGPFAAPKVPADVTQGVDRANGSVRYSY
GKHQENWNAHGPAPERTWDETNGSGRDRDGTQISAPLVVPPPLNGILTNAPIG
TKMDLHFSNFSYGLTAFSHPSVYIQQQIWDKELDLHKPRLHITAFVCKNNAP
GQMLVRLGNLTQDYPNGATLSRVYTGTFFWKGLTWRAKLRAKLTWNPVYQNSVE
DNGNSYMSVTWMLPTATNMQSVPLITRPVARTY"
2792..4558
/gene="vp2"
CDS <2792..4558
/gene="vp2"
/function="coat protein"
/codon_start=1
/product="vp2 protein"
/protein_id="CAB46508.1"
/db_xref="GI:5419929"
/db_xref="GOA:P07302"
/db_xref="Swiss-Prot:P07302"
/translation="TMSDGTGPDGNGVHSAARVERAADPGSGGGSGGGVGVSV
TGSYDNOHYRFLGQWVEITATATRLVHLNMPKSENYCRIRVHNTDTSVKNMAKD
DAHQEIWTPSLVDANAGVWLQPSDQWYICNTMSQLNLSLDOEIFNVLKTVEQD
SGQALKIYNNDLTACMVAVDSNNILPYPYPAANSMTETLGFYPMKFTIASPYRYFCV
DRDLSTVYENQEGTIEHNVMGTPKGMNSQFFTIENQOITLLRTGDEFACTGYVFDN
PVKLTHTWOTNRQLGQPLLSTFPPEADTDAGTLTAQGSRRHATQMENWNAHGP
AOCVPCOPNDPEASAGPFAAPKVPADVTQGVDRANGSVRYSGKHQENWNAHGP
YPERVTWDETNGSGRDRDGTQISAPLVVPPPLNGILTNAPIGTKMDLHFSNFSN
YGLTAFSHPSVYIQQQIWDKELDLHKPRLHITAFVCKNNAPGQMLVRLGNLTQD
YDPNGATLSRVYTGTFFWKGLTWRAKLRAKLTWNPVYQNSVEDNGNSYMSVTWML
PTATNMQSVPLITRPVARTY"
4603..4608
/note="polyadenylation signal"
4821..4826
/note="polyadenylation signal"
4878..5087
/note="terminal hairpin region"

Query Match 94.2%; Score 1902.2; DB 14; Length 5087;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 1946; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
Qy 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA 60
Db 262 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA 321
Qy 61 AGTAACCAAGAAAGTGTCTCATTTGTTTAAATAATGAAATGTTCAACTGAATGAAAA 120
Db 322 AGTAACCAAGAAAGTGTCTCATTTGTTTAAATACTGAGATGTTCAACTGAATGAAAA 381
Qy 121 GATATCGGATGGAATAGTTACAAAAAGAGCTGCAGAGGACGAGCTGGAATCTTTACAA 180
Db 382 GATATCGGATGGAATAGTTACAAAAAGAGCTGCAGAGGACGAGCTGGAATCTTTACAA 441
Qy 181 CGAGGAGCGGAAATCTACTTGGGACCAAGCAGGACATGGAATGGGAAAAACACAGTGGAT 240
Db 442 CGAGGAGCGGAAATCTACTTGGGACCAAGCAGGACATGGAATGGGAAATCTACAGTGGAT 501
Qy 241 GAAATGACCAAAAAAGCAAGTATTTCATTTTGGTATCTTTGGTAAAAAATGTTTATTGAA 300
Db 502 GAAATGACCAAAAAAGCAAGTATTTCATTTATGACTCTTTAGTAAAAAATGTTTATTGAA 561
Qy 301 GTGCTTAAACAAAGAAATATATTTCTGCTGATGTTAAATGGTTGTGCAACATGATGG 360
Db 562 GTGCTTAGCACAAAAAATATAGCTCTCTGCTGATGTTTACTTGGTGTGCAAGATGATGG 621
Qy 361 GGAAGAGACCAAGCGTGGCACTGCCATGTAATTTGGAGGAAGAGCTTTAGTCAAGCT 420
Db 622 GGAAGAGACCAAGCGTGGCACTGCCATGTAATTTGGAGGAAGAGCTTTAGTCAAGCT 681
Qy 421 CAAGGGAATGTTGGAGAGAGGCAACTAAATGTTTACTTGGAGCAGATGTTTGGTAAACAGCC 480
Db 682 CAAGGGAATGTTGGAGAGAGGCACTAAATGTTTACTTGGAGCAGATGTTTGGTAAACAGCC 741
Qy 481 TGTAATGTGCAACTAACACAGCTGAAAGAAATTAATAAGAGAAATAGCAGAGACAAT 540

Db 742 TGTAATGTGAGCTAACACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAACAGT 801
Qy 541 GAGTGGGTACTCTACTTACTTACTTAAAGCATAAGCAACCAAAAGAGACTATACCAAGTGT 600
Db 802 GAGTGGGTACTTACTTACTTAAACATAAGCAACCAAAAGAGACTATACCAAGTGT 861
Qy 601 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAATAAGAAAGAAATTAAGCAGTGT 660
Db 862 GTTCTTTTGGAAATATGATGCTTACTTACTTAACTAATAAGAAAGAAATTAAGCAGTGT 921
Qy 661 CCACCAAGAGCGGAGCTATTTCTTAGCAGTGTACTCTGTGCGGAAAGAAATTAAGCAGTGT 720
Db 922 CCGCCAAGGAGCGGAGCTATTTCTAAGCAGTGTACTCTGTGCGGAAAGAAATTAAGCAGTGT 981
Qy 721 AAAGAGGCGGAGCGGAGCTATTTCTAAGCAGTGTACTCTGTGCGGAAAGAAATTAAGCAGTGT 780
Db 982 AAAGAGGCGGAGCGGAGCTATTTCTAAGCAGTGTACTCTGTGCGGAAAGAAATTAAGCAGTGT 1041
Qy 781 GTTGAACACCAAGTAACTGCGGAGGAACTAAGCGGCGGAGAAATTTCAAACTAAAAA 840
Db 1042 GTTGAACACCAAGTAACTGCGGAGGAACTAAGCGGCGGAGAAATTTCAAACTAAAAA 1101
Qy 841 GAAGTTCTATTAACTACACTTAAGAGCTGTGCGGAGGAAATTTCAAACTAAAAA 900
Db 1102 GAGGTTCTATTAACTACACTTAAGAGCTGTGCGGAGGAAATTTCAAACTAAAAA 1161
Qy 901 GACTGGATGATGATGACGAGCAGTGTACATTTGAATGATGCTCAACGAGGTGGAGAA 960
Db 1162 GACTGGATGATGATGACGAGCAGTGTACATTTGAATGATGCTCAACGAGGTGGAGAA 1221
Qy 961 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAAGCTAGCGGAGGAAACCAAGCA 1020
Db 1222 AACCTGCTGAAAAATACGCTAGAGATTTGTACGCTAAGCTAGCGGAGGAAACCAAGCA 1281
Qy 1021 TTTGACTTAAATTTAGAAAGCTGAAACGAGCAAACTAAGCAATTTTCACTGCTGAC 1080
Db 1282 TTTGACTTAAATTTAGAAAGCTGAAACGAGCAAACTAAGCAATTTTCACTGCTGAC 1341
Qy 1081 ACAAGAGCTGCGAGAAATTTGCTTTTATGCTGGAATCTATGTTAAAGTTTGCATGCT 1140
Db 1342 ACAAGAGCTGCGAGAAATTTGCTTTTATGCTGGAATCTATGTTAAAGTTTGCATGCT 1401
Qy 1141 ATTTGCTGTTTTTAAACAGAGAGGAGGAAAGAAATTAAGCTTTTATTTATGAGCA 1200
Db 1402 ATTTGCTGTTTTTAAACAGAGAGGAGGAAAGAAATTAAGCTTTTATTTATGAGCA 1461
Qy 1201 GCCAGCAGGCAAACTATTATTGCAAGCCTAGCAAGCAGTTGGCATGTTGCT 1260
Db 1462 GCCAGCAGGCAAACTATTATTGCAAGCCTAGCAAGCAGTTGGCATGTTGCT 1521
Qy 1261 TGCTATAATGCGCAATGTAATTTTCCATTTAATGACTGTACCAAGCAAGCAATTTGAT 1320
Db 1522 TGCTATAATGCGCAATGTAATTTTCCATTTAATGACTGTACCAAGCAAGCAATTTGAT 1581
Qy 1321 TGGGTAGAGAAAGCTGGTAACTTTTGGACAGCAAGTAAACAGTTTAAAGCAATTTGCTCT 1380
Db 1582 TGGGTAGAGAAAGCTGGTAACTTTTGGACAGCAAGTAAACAGTTTAAAGCAATTTGCTCT 1641
Qy 1381 GGTCAAACTATTGCGCAATGTAATTTTAAAGGAAAGGAGGAGGAAAGCAAGTAAACCA 1440
Db 1642 GGTCAAACTATTGCGCAATGTAATTTTAAAGGAAAGGAGGAGGAAAGCAAGTAAACCA 1701
Qy 1441 GTCATCATGACCAAAATGAGAACTATTACAGTGGTGTGAGTGTGCGGAGGAAAGCA 1500
Db 1702 GTCATCATGACCAAAATGAGAACTATTACAGTGGTGTGAGTGTGCGGAGGAGCA 1761
Qy 1501 GAACACACTCAACCAATGAGAGAGAGATGCTTAACTTAACTTAACTTAACTTAACTTAA 1560
Db 1762 GAACACACTCAACCAATGAGAGAGAGATGCTTAACTTAACTTAACTTAACTTAACTTAA 1821
Qy 1561 GGTGACTTTGTTGCTTACAAAATGATGCGCCATGTTTGTGCTGTTGTTGTTGTTGTTG 1620

Db 1822 GGTGACTTTGTTGCTTACCAAGAAATGATGCGCCCATGATTTGTGCTGTTGTTGTTG 1881
Qy 1621 AATGGTTACCAATCTACCATGCAAGCTACTGTGCTTAAATGGGCAAAAGTTCTGATTGG 1680
Db 1882 AATGGTTACCAATCTACCATGCAAGCTACTGTGCTTAAATGGGCAAAAGTTCTGATTGG 1941
Qy 1681 TCAGAAAACTGGGCGGAGCAAAAGTGGCAACTCTTATATAATTTACTAGTTGCGGACGC 1740
Db 1942 TCAGAAAACTGGGCGGAGCAAAAGTGGCAACTCTTATATAATTTACTAGTTGCGGACGC 2001
Qy 1741 TCACCAATTCAGCACACGAAAGTGGCTCTCAGCCAGAACTATGCACTTACTTCCACTT 1800
Db 2002 TCACCAATTCAGCACACGAAAGTGGCTCTCAGCCAGAACTATGCACTTACTTCCACTT 2061
Qy 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGCAGCAGCACTTCTCTGTTGCG 1860
Db 2062 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGCAGCAGCACTTCTCTGTTGCG 2121
Qy 1861 GGCACTGCGAGAAACCCAGAACACTGGGGAAGCTGTTCCAAAGCCTGCCAAGATGTCAA 1920
Db 2122 GGCACTGCGAGAAACCCAGAACACTGGGGAAGCTGTTCCAAAGCCTGCCAAGATGTCAA 2181
Qy 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTCGTGGGAAACCG 1980
Db 2182 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTCGTGGGAAACCG 2241
Qy 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTTGGACTAA 2019
Db 2242 TTGAAGAGAGACTTCAGCGAGCGCTGAACTTTGGACTAA 2280

RESULT 11

MPUI2469
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

MPUI2469
5144 bp DNA linear VRL 24-JAN-1995
Mouse parvovirus 1 complete genome, nonstructural protein 1 (NS1)
gene, complete cds and capsid protein (VP1) gene, complete cds.
U12469
U12469.1 GI:525325
Mouse parvovirus 1
Mouse parvovirus 1
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
1 (bases 1 to 5144)
Ball-Goodrich, L.J. and Johnson, E.
Molecular characterization of a newly recognized mouse parvovirus
J. Virol. 68 (10), 6476-6486 (1994)
94365951
8083985
2 (bases 1 to 5144)
Ball-Goodrich, L.J.
Direct Submission
Submitted (20-JUL-1994) Lisa J. Ball-Goodrich, Section of
Comparative Medicine, Yale University School of Medicine, New
Haven, CT 06520-8016, USA

FEATURES

1..5144
/organism="Mouse parvovirus 1"
/mol_type="genomic DNA"
/db_xref="taxon:35340"
/clone="pRVXB36; pLVB/R13-6; pRTBg/H3-1"
265..2283
/gene="NS1"
265..2283
/gene="NS1"
/codon_start=1
/product="nonstructural protein 1"
/protein_id="AAA61405.1"
/db_xref="GI:525326"
gene
CDS

OTKKEYSIKITTLKELVHKVTSPEDEMMQPDSDSYIEWMAQPGGNNLKNLTBICITL
ARTKFDLLEKAEKTSKUNFSPDTRICKIFAPFGWNTIKVCHCAICVLNRQGGKR
NTVLFPGATSGKSIIAQAIAQAVGNVYNAANVPFNDCTKNKLIWVEAGNFGQ
OWNQPAICSGQIIRIDQKGGSKQIEPTVIMTNENITVVRIGCESEPEHTQPIRD
RMLNHLHTLPGDFGLVDKNWPMICAWLVKNGYOSTWASYCAKWKGVKPDWENWAE
PKVPTINSLSGARSPTTPKSPPLSONVALTPLASDLEDLALPEWSTENTPVAGTAE
TONTGAGSKACODGQSLSPWBSIEDLRACFGAEPKXDFSEPLND
join(2290..2320,2403..4561)
/gene="Vp1"
CDS
join(2290..2320,2403..4561)
/gene="Vp1"
/codon_start=1
/product="capsid protein"
/protein_id="AAA61406.1"
/db_xref="GI:525327"

AVDKYIKSGKNPFLYFSAADOFIDOTKADKGGKGVHVFRTKRAFAPRLASSEPE
GTSVGIAGKRTKPPAHIFINOARAKKRSLAAQOORTOTMSDGAOPDSSGSAVQSA
RVERAADGPGSGSGGGGVSTGSDYNQTHYRFLSDGWVEITATYSTRVHLNWP
KSENYCRVRHNTNDRKAGNMAKDAHQIWIWPLSDSNWGVWFPDSDWQFICNN
MSHVNLSLDQELFNVIKTVTEQNTGAEAVKIYNNDLTASMMVALDSNNILPYTPAT
DNQETLGFYKPEKTFMSPRYFYFNCDSLSVTYDQTSIVDTMANASGLSQFFIE
NTORIQLLRGTDFATGTYFTEPEIKLSHTWOSNRQLGPOPOITDLPADNENATLV
TRGDRSISITQISGNDVTEATRVRPAQVGFPCOPHDNFEFSRAGPKVPVVPVANVTGN
EHDANGSLRYTDKQGDQWGSNNKERTWDALSYDSGRWADRCFINATPFTSPPAL
NNLITNSDPIGNKTAIHYQNVFNSYGLTAFPHAPPIYPOGQIWDKELDEHKPLRHA
QAPFVCKNNAPOGLLVRAPNLTDQYDPSNLSLRYTYGTFPMKGLKTKAKLRPNA
TWNPVYQVSAQYQENENYMSIHKWLPATATGNMQSIPLLSRPARNTY"

ORIGIN

Query Match		94.1%; Score 1899; DB 14; Length 5144;
Best Local Similarity		96.3%; Pred. No. 0;
Matches 1944; Conservative		0; Mismatches 75; Indels 0; Gaps 0;
Qy	1	ATGGCTGGAAATCCTTACTCTGATGAAGTTTTCGGAGCAACCACTGGTTTAAAGGAAAA 60
Db	265	ATGGCTGGAAATCCTTACTCTGATGAAGTTTTCGGAGCAACCACTGGTTTAAAGGAAAA 324
Qy	61	AGTAACGAGGAAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAAATGGAATA 120
Db	325	AGTAACGAGGAAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAAATGGAATA 384
Qy	121	GATATCGEATGGAATAGTTACAAAAAGAGCTCGAGGAGACAGCTGAAATCTTTACAA 180
Db	385	GATATCGEATGGAATAGTTACAAAAAGAGCTCGAGGAGACAGCTGAAATCTTTACAA 444
Qy	181	CGAGGAGCGAATACTACTTGGGACCAAGAGGAGCATGGAATGGGAAACCAAGTGGAT 240
Db	445	CGAGGAGCGAATACTACTTGGGACCAAGAGGAGCATGGAATGGGAAATCTACAGTGGAT 504
Qy	241	GAAATGACCAAAAAGCAAGTATTCATTTTGTATCTTTGTTTAAAAATGTTTTATTTGAA 300
Db	505	GAAATGACCAAAAAGCAAGTATTCATTTATGACTCTTTAGTTAAAAAATGTTTGTGAA 564
Qy	301	GTGCTTAAACAAAGAAATATATTTCTCTGGTGAATGTTAAATGGTTGTGCAACATGAATGG 360
Db	565	GTGCTTAAACAAAGAAATATAGCTCTCTGCTGATGTTACTTGGTTTGTGCAACATGAATGG 624
Qy	361	GGAAAGACCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAGGACTTTTGTCTCAAGCT 420
Db	625	GGAAAGACCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAGGACTTTTGTCTCAAGCT 684
Qy	421	CAAGGAAATGGTGGAGAGGCACTAAATGTTTTTACTTGGAGCAGATGGTTGTTAAACAGCC 480
Db	685	CAAGGAAATGGTGGAGAGGCACTAAATGTTTTTACTTGGAGCAGATGGTTGTTAAACAGCC 744
Qy	481	TGTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGAACAAT 540
Db	745	TGTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGAACAAT 804
Qy	541	GAGTGGGTTACTTACTTACTTATAGCATAAGCAACCAACCAAGACTATACCAAGTGT 600
Db	805	GAGTGGGTTACTTACTTACTTATAGCATAAGCAACCAACCAAGACTATACCAAGTGT 864

Qy	601	GTTCCTTTTGGAAAAATGATGTTGCTTACTTATTTTTTAACTAAAAAGAAAAATAGCACTAGT 660
Db	865	GTTCCTTTTGGAAAAATGATGTTGCTTACTTATTTTTTAACTAAAAAGAAAAATAGCACTAGT 924
Qy	661	CCACCAAGAGACGGAGGCTATTTCTTACGAGTACTCTGGCTGGAAAACTTAACCTTTTAA 720
Db	925	CCGCCAAGAGACGGAGGCTATTTCTTAAAGCAGTACTCTGGCTGGAAAACTTAACCTTTTAA 984
Qy	721	AAAGAAAGGGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAAAAG 780
Db	985	AAAGAAAGGGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAAAAG 1044
Qy	781	GTGAAAAACACAGTAAACCACTGCGGAGGAAAATAAGCGGCGGAGAAATTTCAAACCTAAAAA 840
Db	1045	GTGAAAAACACAGTAAACCACTGCGGAGGAAAATAAGCGGCGGAGAAATTTCAAACCTAAAAA 1104
Qy	841	GAAAGTTTCTATTAAAACTACCTTAAAGAGCTGGTGCATAAAAAGAGTAACCTCACCAGAG 900
Db	1105	GAGGTTTCTATTAAAAACCACTTAAAGAGCTGGTGCATAAAAAGAGTAACCTCACCAGAA 1164
Qy	901	GACTGATGATGATGACGAGCAGACAGTTACATTTGAAATGATGCTCAACACGAGTGGAGAA 960
Db	1165	GACTGATGATGATGACGAGCAGACAGTTACATTTGAAATGATGCTCAACACGAGTGGAGAA 1224
Qy	961	AACCTCTGAAAAAATACGCTAGAGATTGTACACTAACTCTACCGAGAACCAAAAAACGCA 1020
Db	1225	AACCTCTGAAAAAATACGCTAGAGATTGTACACTAACTCTACCGAGAACCAAAAAACGCA 1284
Qy	1021	TTTGACTTAAATTTAGAAAAAGCTGAAACCGAGCAAACTAACCACTTTTCACTGCCTGAC 1080
Db	1285	TTTGACTTAAATTTAGAAAAAGCTGAAACCGAGCAAACTAACCACTTTTCACTGCCTGAC 1344
Qy	1081	ACAAGAGCCTGCAAGAAATTTTGGCTTTTCATGGCTGGAACCTATCTTAAAGTTTGCCTGCT 1140
Db	1345	ACAAGAGCCTGCAAGAAATTTTGGCTTTTCATGGCTGGAACCTATCTTAAAGTTTGCCTGCT 1404
Qy	1141	ATTGCTGTGTTTTTAAACAGAGAGGAGGAAAAAATACTGTTTTTATTTTCATGGACCA 1200
Db	1405	ATTGCTGTGTTTTTAAACAGAGAGGAGGAAAAAATACTGTTTTTATTTTCATGGACCA 1464
Qy	1201	GCAGACAGGCAAAATCTATTATTCAGAACCATAGCAGCAAGCTTGGCAATGTTGGT 1260
Db	1465	GCAGACAGGCAAAATCTATTATTCAGAACCATAGCAGCAAGCTTGGCAATGTTGGT 1324
Qy	1261	TGCTATAATGACGCAATGTAAACTTTTCAATTTAAATGACTGTACCAACCAAGAACTTGATT 1320
Db	1525	TGCTATAATGACGCAATGTGAACTTTTCAATTTAAATGACTGTACCAACCAAGAACTTGATT 1384
Qy	1321	TGGGTAGAAAGAGCTGGTAACCTTTTGGACAGCAAGTAAACAGTTTAAAGCCTTTGCTCT 1380
Db	1585	TGGGTAGAAAGAGCTGGTAACCTTTTGGACCAACAGTAAACAGTTTAAAGCCTTTGCTCT 1644
Qy	1381	GGTCAAACTATTTCGCAATTCATCAAAAGGAAAGGAGGAGCAAAACAGATTGAAACCAACCA 1440
Db	1645	GGTCAAACTATTTCGCAATTCATCAAAAGGAAAGGAGGAGCAAAACAGATTGAAACCAACCA 1704
Qy	1441	GTCAATCATGACCAAAATGAGAACATTTACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA 1500
Db	1705	GTCAATCATGACCAAAATGAGAACATTTACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA 1764
Qy	1501	GAAACACATCAACCAATCAGAGACAGAAATGCTTTAACTTATCATCTAAACATCTTGCCT 1560
Db	1765	GAAACACATCAACCAATCAGAGACAGAAATGCTTTAACTTATCATCTAAACATCTTGCCT 1824
Qy	1561	GGTGACTTTGGTTGGTTGACAAAAATGAAATGGCCCCATGATTTGCTGCTGGTTGGTAAAG 1620
Db	1825	GGTGACTTTGGTTGGTTGACAAAAATGAAATGGCCCCATGATTTGCTGCTGGTTGGTAAAG 1684
Qy	1621	AATGGTTTACCAATCTACCACTGCAAGCTACTGTGCTTAAATGGGGCAAAAGTTCTCTGATTGG 1680
Db	1885	AATGGTTTACCAATCTACCACTGCAAGCTACTGTGCTTAAATGGGGCAAAAGTTCTCTGATTGG 1944


```
QY 1681 TCAGAAACCTGGGCGGAGCCAAAGGTGCCAACTCCTATAAATTTACTAGTTCGGCAGC 1740
Db 1945 ACGGAAACCTGGGCGGAGCCGAGGTGCCACTCCTATAAATTTACTAGTTCGGCAGC 2004
QY 1741 TCACCATTCAGACACACGAAAGTACCGCTCTCAGCAGAACTATGCACTAATCCACTT 1800
Db 2005 TCACCATTCAGACACACGAAAGTACCGCTCTCAGCAGAACTATGCACTAATCCACTT 2064
QY 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGGAGCACACCAATACTCTCTGTTGCG 1860
Db 2065 GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGGAGCACACCAATACTCTCTGTTGCG 2124
QY 1861 GGCACTCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920
Db 2125 GGCACTCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 2184
QY 1921 CTGAGCCCACTTGGTCAGAGATCGAGGAGATTTGAGACGCTGCTTCGCTGCGGAACCG 1980
Db 2185 CTGAGCCCACTTGGTCAGAGATCGAGGAGATTTGAGACGCTGCTTCGCTGCGGAACCG 2244
QY 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2019
Db 2245 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2283

RESULT 12
LOCUS MOU34254
DEFINITION Mouse parvovirus 1c DNA.
ACCESSION U34254
VERSION U34254.1 GI:1464794
KEYWORDS .
SOURCE Mouse parvovirus 1c
ORGANISM Mouse parvovirus 1c
REFERENCE 1 (bases 1 to 4764)
AUTHORS Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L., Franklin,C.L., Hook,R.R. Jr. and Riley,L.K.
TITLE Molecular characterization of newly recognized rodent parvoviruses
JOURNAL J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
MEDLINE 96201434
PUBMED 8609486
REFERENCE 2 (bases 1 to 4764)
AUTHORS Besselsen,D.G.
TITLE Direct Submission
JOURNAL Submitted (17-AUG-1995) David G. Besselsen, Department of Veterinary Pathology, University of Missouri-Columbia, W213 Veterinary Medicine Building, Columbia, MO 65211, USA

FEATURES
source Location/Qualifiers
1..4764
/organism="Mouse parvovirus 1c"
/mol_type="genomic DNA"
/specific_host="Mus musculus"
/db_xref="taxon:42842"

ORIGIN
Query Match 92.8%; Score 1873; DB 14; Length 4764;
Best Local Similarity 95.4%; Pred. No. 0;
Matches 1927; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAAA 60
Db 122 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAAA 181
QY 61 AGTAACGAGGAAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGAAAA 120
Db 182 AGTAACGAGGAAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGAAAA 241
QY 121 GATATCGGATGGAATGTTACAAAAAGAGCTCGAGGAGCAAGCTGGAATCTTTACAA 180
Db 242 GATATGGAATGGAATGTTACAAAAAGAGCTCGAGGAGCAAGCTGGAATCTTTACAA 301
QY 181 CGAGGAGCGGAACTACTTGGGACCAAGCGAGCATGGAATGGGAAACCACTGGAT 240
```

```
Db 302 CGAGGAGCGGAACTACTTGGGACCAAGCGAGCATGGAATGGGAATCTACAGTGGAT 361
QY 241 GAAATGACCAAAAGCAAGTATTCATTTTGTCTTTGTTTAAAAAATGTTTATTTGAA 300
Db 362 GAAGTACCAAAAGCAAGTATTCATTTTGTATGACTCTTTAGTTAAAAAATGTTTGTGAA 421
QY 301 GTGCTTAAACAAAGAAATATATTTCTCGTGTATGTTAAATTTGGTTTGTCAACATGAATGG 360
Db 422 GTGCTTAAACAAAGAAATATCTCTCGTGTATGTTAAATTTGGTTTGTGAGCATGAATGG 481
QY 361 GGAAGAACAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAGCACTTTAGTCAAGCT 420
Db 482 GGAAGAGACAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAGCACTTTAGTCAAGCT 541
QY 421 CRAAGGAATGTGGAGAGGCAACTAAATGTTTACTGAGCGAGATGTTGTTGTAAACGCC 480
Db 542 CRAAGGAATGTGGAGAGGCAAGCTAAGTGTCTTACTGAGCAGATGGTTGGTAAACGCC 601
QY 481 TGTAAATGTCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAACAAT 540
Db 602 TGTAAATGTACAGCTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAACAAT 661
QY 541 GAGTGGTTTACTCTTACTTACTTAAAGCAATAGCAAAACCAAAAGCACTATACCAAGTGT 600
Db 662 GAATGGTTTACTCTTACTTACTTAAAGCAATAGCAAAACCAAAAGCACTATACCAAGTGT 721
QY 601 GTTCTTTTGGAAACATGATTCCTTACTATTTTAACTTAAAGAAATAAGCACTACT 660
Db 722 GTTCTTTTGGCAACATGATTCCTTACTATTTTAAACCAAGAAATAAGCACTACT 781
QY 661 CCACCAAGAGACGAGGCTTATTTCTTAGCAGTACTCTGGCTGGAAAACTAATCTTTTAA 720
Db 782 CCGCCAAAGGACGAGGCTTATTTCTTAGTACTGCTCTGGCTGGAAAACTAATCTTTTAA 841
QY 721 AAGAAGGCGAGCGGCATCTAGTGACAACTATACACTGATGACNTGCGGCGCAGAAACG 780
Db 842 AAGAAGGCGAGCGGCATCTAGTGACAACTATACACTGATGACNTGCGGCGCAGAAACG 901
QY 781 GTTGAACCAACAGTAACTACCTCGCGAGGAACTAAGCGCGGAGAACTTCAAACTAAAAA 840
Db 902 GTTGAACCAACAGTAACTACCTCGCGAGGAACTAAGCGCGGAGAACTTCAAACTAAAAA 961
QY 841 GAAGTTTCTTATTTAAAACTACACTTAAAGAGCTGGTGCAATAAAGAGTAACTCACCAGAG 900
Db 962 GAGGTTTCTTATTTAAAAACCACTTAAAGAGCTAGTGCAATAAGAGATTAACCTCACAGAA 1021
QY 901 GACTGATGATGATGCGAGCAGACAGTTACATTGAAATGATGCTCAACAGCTGGAGAA 960
Db 1022 GACTGATGATGATGCGAGCAGACAGTTATATTGAAATGATGCTCAACAGCTGGAGAA 1081
QY 961 AACCTGCTGAAAAATACGCTAGAGATTTGTACATACTCTAGCCAGAAACCAAAACAGCA 1020
Db 1082 AACCTGCTGAAAAATACGCTAGAGATCTGTACACTTAACTCTAGCTAGAAACCAAAACAGCA 1141
QY 1021 TTTGACTTAAATTTAGAAAAAGCTGAAACCAAGCAAACTAACCAACTTTTCACTGCTGAC 1080
Db 1142 TTTGACTTAAATTTAGAAAAAGCTGAAACCAAGCAAACTAACCAACTTTTCACTGCTGAC 1201
QY 1081 ACAAGAGCTGAGAAATTTTGTCTTTTCTGCTCGGAACTATGTTAAAGTTTGGCAATGCT 1140
Db 1202 ACAAGAGCTGAGAAATTTTGTCTTTTCTGCTCGGAACTATGTTAAAGTTTGGCAATGCT 1261
QY 1141 ATTTGCTGTGTTTTTAAACAGACAGAGGCAAAAGAAATACCTGTTTATTTTCATGGACCA 1200
Db 1262 ATTTGCTGTGTTTTTAAACAGACAGAGGCAAAAGAAATACCTGTTTATTTTCATGGACCA 1321
QY 1201 GCCAGACAGGCAAAATCTATTTTCAAGCAAGCTATGACCAAGAGCTGGCAATGTTGCT 1260
Db 1322 GCCAGTACAGGCAAAATCTATTTTCAAGCAAGCTATGACCAAGAGCTGGCAATGTTGCT 1381
QY 1261 TGCTATATGACGCAATCTTAACTTTTCAATTTATGACTGTACCAACCAAGAACTTGATT 1320
```



```
Db 1382 TGTTATATGACGAATGTGAACCTTTCCATTCAATGACTGTACCAACAGAACTTAATT 1441
Qy 1321 TGGGTAGAAGAGCTGGTAACTTTTGGACAGACAGTAAACAGATTTAAAGCAATTTGCTCT 1380
Db 1442 TGGGTGGAAGAGCTGGTAACTTTTGGACACAAAGTAAACAGATTTAAAGCAATTTGCTCT 1501
Qy 1381 GGTCAAACTATTCCGATTTGATCAAAAGGAAAGGAGGAGCAAAACAGATTTGAACCAACACCA 1440
Db 1502 GGTCAAACTATTCCGATTTGATCAAAAGGAAAGGAGGAGGAGGAGGAGGATTTGAACCAACACCA 1561
Qy 1441 GTCATCATGACCAACAAATGAGAACATTTACAGTGGTTCAGATAGGCTGCGAAGAAAGACCA 1500
Db 1562 GTCATCATGACCAACAAATGAGAACATTTACAGTGGTTCAGATAGGCTGCGAAGAAAGACCA 1621
Qy 1501 GAACACACTCAACCAATGAGAGACAGATGCTTAAACATTCATCAACATCAACATCAACATCA 1560
Db 1622 GAACACACTCAACCAATGAGAGACAGATGCTTAAACATTCATCAACATCAACATCAACATCA 1681
Qy 1561 GGTGACTTTGGTTGGTTGACAAACATGANTGGCCCATGATTTGTGCTTGGTTGGTAAAG 1620
Db 1682 GGTGACTTTGGTTGGTTGACAAACATGANTGGCCCATGATTTGTGCTTGGTTGGTAAAG 1741
Qy 1621 AATGGTTTACCAATCTACCTGGAAGCTACTGTGCTAAATGGGGCAAAAGTTCTGATTTGG 1680
Db 1742 AATGGTTTACCAATCTACCTGGAAGCTACTGTGCTAAATGGGGCAAAAGTTCTGATTTGG 1801
Qy 1681 TCAGAAAACCTGGCGGAGCCAAAGGTGCCAACTCTTATATAATTTACTAGTTCGGCACGC 1740
Db 1802 ACGAAAACCTGGCGGAGCCAAAGGTGCCAACTCTTATATAATTTACTAGTTCGGCACGC 1861
Qy 1741 TCACCAATTACGACACCGAAGAGTACGCTCTCAGCAGAACATGCACTTAACCTCACTT 1800
Db 1862 TCACCAATTACGACACCGAAGAGTACGCTCTCAGCAGAACATGCACTTAACCTCACTT 1921
Qy 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGAGACCTTGAGACACCAATATCTCTGTTGG 1860
Db 1922 GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGAGACCTTGAGACACCAATATCTCTGTTGG 1981
Qy 1861 GGCACTGCGAAGAACCCAGAACACTGGGGAGCTGGTTTCCAAAGCTTCCCAAGATGTCAA 1920
Db 1982 GGCACTGCGAAGAACCCAGAACACTGGGGAGCTGGTTTCCAAAGCTTCCCAAGATGTCAA 2041
Qy 1921 CTGAGCCCACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGCTTCGGTGGGAAACCG 1980
Db 2042 CTGAGCCCACTTGGTCAGAGATCGAGGAGGATTTGAAAGCGTGTTCGGTGGGAAACCG 2101
Qy 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2019
Db 2102 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2140

RESULT 13
HOU34255
LOCUS Hamster parvovirus DNA.
DEFINITION Hamster parvovirus DNA.
ACCESSION U34255
VERSION U34255.1 GI:1464792
KEYWORDS
SOURCE
ORGANISM Hamster parvovirus
VIRUSES; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE
1 (bases 1 to 4773)
AUTHORS Besselsen, D.G., Pintel, D.J., Purdy, G.A., Besch-Williford, C.L., Franklin, C.L., Hook, R.R. Jr. and Riley, L.K.
Molecular characterization of newly recognized rodent parvoviruses
J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
JOURNAL
MEDLINE 96201434
PUBMED 8609486
REFERENCE
2 (bases 1 to 4773)
AUTHORS Besselsen, D.G.
Direct Submission
TITLE Submitted (17-AUG-1995) David G. Besselsen, Department of
JOURNAL Veterinary Pathology, University of Missouri-Columbia, W213
```

```
FEATURES
    source
    Location/Qualifiers
        /organism="Hamster parvovirus"
        /mol_type="genomic DNA"
        /specific_host="Mesocricetus auratus"
        /db_xref="taxon:42843"

ORIGIN
    Query Match ^ 92.2%; Score 1862.2; DB 14; Length 4773;
    Best Local Similarity 95.1%; Pred. No. 0;
    Matches 1921; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

Qy 1 ATGGCTGGAAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA 60
Db 122 ATGGCTGGAAAATGCTTACTCTGATGAAGTTTGGGAAACCAACCAACTGGTTAAAGGAAA 181
Qy 61 AGTAAACAGAGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGGAATA 120
Db 182 AGTAAACAGAGAGTGTCTCATTTGTTTAAAAATGAAAGATGTTTCAGCTCAATGGAATA 241
Qy 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGAGGAGCAGCTGAAATCTTTACAA 180
Db 242 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGAGGAGAGCTGAAATCTTTACAA 301
Qy 181 CGAGGAGCGGAAAACCTACTCTGGGACCAAGCGAGGACATGGAATGGGAAACCAAGTGGAT 240
Db 302 CGAGGAGCGGAAAACCTACTCTGGGACCAAGCGAGGACATGGAATGGGAAATCTTCAGTGGAT 361
Qy 241 GAAATGACCAAAAGCAAGTATTCATTTGTTTGTATTTCTTTGGTTAAAAAATGTTTATTGAA 300
Db 362 GAAATGACCAAAAGCAAGTATTCATTTGTTTGTATTTCTTTAGTTAAAAAATGTTTATTGAA 421
Qy 301 GTGCTTAAACACAAAGAAATATATTTCTCTGGTGAATGTTAAATGTTGTTGCAACATGAATGG 360
Db 422 GTGCTGAGTACAAAGAAACATAGCACCTAGTGAATGTTGTTGTTGTTGACGATGAATGG 481
Qy 361 GGAAGAGACCAAGGCTGGCACTGCCATGTAATTTGGAGGAAAGCACTTTAGTCAAGCT 420
Db 482 GGAAGAGACCAAGGCTGGCACTGTCATGTAATTTGGAGGCAAGGACTTTAGCCAAGCT 541
Qy 421 CAAGGAAATGCTGGAGAGGCAACTAAATGTTTACTGAGCAGATGTTGTTGAACAGCC 480
Db 542 CAAGGAAATGCTGGAGAGGCAAGTAAATGTTTACTGGAGCAGATGTTGTTGAACAGCC 601
Qy 481 TGTAAATGTGCAACTAACACAGCTGAAAGAAATTTAAACTAAGAGAAATAGCAGAAAGCAAT 540
Db 602 TGTAGTGTGAGCTATTACAGCTGAAAGAAATTTAAGCTGAGAGATAGCGGAAGACCA 661
Qy 541 GAGTGGTTTACTCTACTTACTTATAGCATATAAGCAACCAACCAACCAATATACCAAGTGT 600
Db 662 GAATGGTCACTTTGCTTACTTATAGCATATAAGCAACCAACCAACCAATATACCAAGTGT 721
Qy 601 GTTCTTTTGGAAACATGATGCTTACTATTTTTAACTAAAAAGAAATTAAGCACTAGT 660
Db 722 GTTGTCTTTGGAATATAGTCTTACTACTTTTATCCAAAGAAAGAAATATATGCAACAGT 781
Qy 661 CCACCAAGAGAGCGAGGCTATTCTTACAGTGAATGCTGCTGGTGGAAAACTTAATCTTTTA 720
Db 782 CCACCAAGAGAGCGAGGCTATTCTTACAGTGAATGCTGCTGGTGGAAAACTTAATCTTTTA 841
Qy 721 AAAGAGGCGAGCGGCATCTAGTGAGCAAACTATACATCTGATGACATGCGGCCAGAAACG 780
Db 842 AAAGAGGCGAGCGGCATCTAGTGAGCAAACTATACATCTGATGACATGCGGCCAGAAACG 901
Qy 781 GTTGAACACACAGTAAACCACTGCGGAGGAAACTAAGCGGCGGAGATTTCAAACTAAAAAA 840
Db 902 GTTGAACACACAGTAAACCACTGCGGAGGAAACTAAGCGGCGGAGATTTCAAACTAAAAAA 961
Qy 841 GAAGTTTCTTATTAACACTACACTTTAAAGAGCTGGTGCATATAAGAGAGTAACTCACCAGAG 900
Db 962 GAGGTCTCTATTAAAAACCACTTAAAGAGCTGGTGCATATAAGAGAGTAACTCACCAGAA 1021
```

Veterinary Medicine Building, Columbia, MO 65211, USA

```
Qy 901 GACTGGATGATGATGACGAGCAGACAGTTACATTGAAATGATGCTCAACAGGTGGAGAA 960
Db 1022 GACTGGATGATGATGACGAGCAGACAGTTACATTGAAATGATGCTCAACAGGTGGAGAA 1081
Qy 961 AACCTGCTGAAATATACGCTAGAGATTGTACACTAACTCTAGCCAGAACCAAAACAGCA 1020
Db 1082 AACCTGCTGAAATATACGCTAGAGATTGTACACTAACTCTAGCCAGAACCAAAACAGCA 1141
Qy 1021 TTTGACTTAATTTTAGAAAAAGCTGAAACACGAGAACTAACCAACTTTTTCACCTGCTGAC 1080
Db 1142 TTTGACTTAATTTTAGAAAAAGCTGAAACACGAGAACTAACCAACTTTTTCACCTGCTGAC 1201
Qy 1081 ACAAGACCTGCAAAATTTTGTCTTTTCATGGCTGGAATCATGTTAAAGTTTCCCATGCT 1140
Db 1202 ACAAGACCTGCAAAATTTTGTCTTTTCATGGCTGGAATCATGTTAAAGTTTCCCATGCT 1261
Qy 1141 ATTGCTGTGTTTAAACAGACAGAGGCAAAAGAAATACTGTTTATTTTATGACCA 1200
Db 1262 ATTGCTGTGTTTAAACAGACAGAGGCAAAAGAAATACTGTTTATTTTATGACCA 1321
Qy 1201 GCCAGACAGCGCAAACTATTTATGCAACAGCCATAGCAACAGCTTGGCAATGTTGGT 1260
Db 1322 GCCAGTACAGCAAAATCCATTTATGCAAGCCATAGCAACAGCTTGGTAAATGTTGGT 1381
Qy 1261 TGCTATTAATGACGCAATGTAAACTTTTCCATTTAATGACTGTACCAACAGAACTTTGAT 1320
Db 1382 TGCTATTAATGACGCAATGTAAACTTTTCCATTTAATGACTGTACCAACAGAACTTTGAT 1441
Qy 1321 TGGGTAGAAAGCTGGTAACTTTTGACACAGCAAGTAAACAGTTTAAAGCCATTGCTCT 1380
Db 1442 TGGGTAGAAAGCTGGTAACTTTTGACACAGCAAGTAAACAGTTTAAAGCCATTGCTCT 1501
Qy 1381 GGTCAAACTATTGCTATTGATCAAAAGGAAAGGAGCAACAGATTGAAACCAACCA 1440
Db 1502 GGTCAAACTATTGCTATTGATCAAAAGGAAAGGAGCAACAGATTGAAACCAACCA 1561
Qy 1441 GTCATCATGACCAAAATGAGAACTTACAGTGGTCAGATAGGCTGCGAAGAAAGACCA 1500
Db 1562 GTCATCATGACCAAAATGAGAACTTACAGTGGTCAGATAGGCTGCGAAGAAAGACCA 1621
Qy 1501 GAACACACTCAACCAATCAGACAGACAGATGCTTAAACATTTCATCTAACACATCTTGCCT 1560
Db 1622 GAACACACTCAACCAATCAGACAGACAGATGCTTAAACATTTCATCTAACACATCTTGCCT 1681
Qy 1561 GGTGACTTTTGGTTGTTGACAAAAATGAAATGGCCCATGATTTGCTGTTGGTTGTTAAAG 1620
Db 1682 GGTGACTTTTGGTTGTTGACAAACATGATGGCCCATGATTTGCTGTTGGTTGTTAAAG 1741
Qy 1621 AATGGTTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCTGATTGG 1680
Db 1742 AATGGTTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCTGATTGG 1801
Qy 1681 TCAGAAACTGGGGGAGCAAAAGGTGCCAACTCTTATATAATTTACTAGTTTGGCACGC 1740
Db 1802 ACGGAAACTGGGGGAGCGAAGGTGCCAACTCTTATATAATTTACTAGTTTGGCACGC 1861
Qy 1741 TCACCAATTCACGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTAATCTCCACTT 1800
Db 1862 TCACCAATTCACGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTAATCTCCACTT 1921
Qy 1801 GCATCGGATCTCGAGGACCTGGCTTTTAGAGCCCTTGAGACACACCAAAATCTCCTGTTGG 1860
Db 1922 GCATCGGATCTCGAGGACCTGGCTTTTAGAGCCCTTGAGACACACCAAAATCTCCTGTTGG 1981
Qy 1861 GGCACCTGCAGAAACCCAGACACTGGGGAGCTGGTTCCAAAGCCTGCCAAGTGTGCA 1920
Db 1982 GGCACCTGCAGAAACCCAGACACTGGGGAGCTGGTTCCAAAGCCTGCCAAGTGTGCA 2041
Qy 1921 CTGAGCCCAACTTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTCGGTGCGAACCG 1980
Db 2042 CTGAGCCCAACTTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTCGGTGCGAACCG 2101
Qy 1981 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019
```

```
Db 2102 TTGAAGAGAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2140
RESULT 14
MVU34256
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .4761
/organism="Mice minute virus"
/mol_type="genomic DNA"
/strain="Cutter"
/specific_host="Mesocricetus auratus"
/db_xref="taxon:10794"
ORIGIN
Query Match 91.9%; Score 1855.8; DB 14; Length 4761;
Best Local Similarity 94.9%; Pred. No. 0;
Matches 1917; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
Qy 1 ATGGCTGGAAATGCTTACTCTGATGAACTTTTGGGAGCAACCACTGGTTAAAGNAAA 60
Db 122 ATGGCTGGAAATGCTTACTCTGATGAGGTTTGGGAGCAACCACTGGTTAAAGNAAA 181
Qy 61 AGTAACACGAGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAA 120
Db 182 AGTAACACGATGTAGTATCTCATTTGTTTAAAAATGAAGATGTTCAATGATGAAAA 241
Qy 121 GATATCGGATGAAATAGTTTACAAAAAGAGCTGCGAGGAGCAGCTGAAATCTTTTACAA 180
Db 242 GATATCGGATGAAATAGTTTACAAAAAGAGCTGCGAGGAGCAGCTGAAATCTTTTACAA 301
Qy 181 CGAGGAGCGGAACTTACTTGGGACCAACGAGGACATGGAATGGGAAACCACTGGAT 240
Db 302 CGAGGAGCGGAACTTACTTGGGACCAACGAGGACATGGAATGGGAACTTTCAGTGGAT 361
Qy 241 GAAATGACCAAAAAAGCAAGTATTCTTTTGTGTTTAAAAATGTTTATTTGAA 300
Db 362 GAACTAACCAACAGCAAGTATTCTATTTTGTCTTTAGTTTAAAAAGTGTATTGAA 421
Qy 301 GTGCTTAAACAAAGAATAATTTCTGTTGATGTTTAAATGTTGTGCAACATGAATGG 360
Db 422 GTGCTAAGTACAAAGAAACATAGCTCTAGTGTATTTAAATGTTGTGTCAGCATGAATGG 481
Qy 361 GGAAGACCAAGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGCACTTTAGTCAAGCT 420
Db 482 GGAAGACCAAGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGCACTTTAGTCAAGCT 541
Qy 421 CAAGGAAATGCTGAGAGGCAACTAAATGTTTACTTGAGCAGATGTTGTTGTAACAGCC 480
Db 542 CAAGGAAATGCTGAGAGGCACTAAATGTTTACTTGAGCAGATGTTGTTGTAACAGCC 601
Qy 481 TGTAAATGTCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT 540
```

Db 602 TGCAGTGTGAGCTATCACAGCGGAAAGAAATTAAGCTGAGAGAAATAGCGGAAGACCA 661
Qy 541 GAGTGGGTACTCTACTTACTTATAAGCATAAGCAAAACCAAAAGAGACTATACCAAGTGT 600
Db 662 GAATGGGTCACTTGTCTTACTTATAAGCATAAGCAAAACCAAAAGAGACTATACCAAGTGT 721
Qy 601 GTTCTTTTGGAAACATGATGCTTACTTATTTTAACTAAAGAAAGAAATTAAGCAGTAGT 660
Db 722 GTTGTCTTGGAAATATGATGTCTTACTTATTTTAAACCAAGAAAGAAATATGACCTAGT 781
Qy 661 CCACCAAGAGACGAGGCTATTTCTTGTAGCAGTACTCTCGCTGGGAAAGCTAACTTTTAA 720
Db 782 CCACCAAGGAGCGAGGCTATTTCTTGTAGCAGTACTCTCGCTGGGAAAGCTAACTTTTAA 841
Qy 721 AAAGAAAGGCGAGGCGCCATCTAGTGAGCAAACTATACCTGATGACATGCGCGCAGAAACG 780
Db 842 AAAGAAAGGCGAAACGCCATCTAGTGAGCAAACTATACCTGATGACATGCGCGCAGAAACG 901
Qy 781 GTTGAACCAACAGTAACCACTGGCGAGGAAACTAAGCGCGGAGAAATTTCAAACTAAAGAA 840
Db 902 GTTGAACCAACAGTAACCACTGGCGAGGAAACTAAGCGCGGAGAAATTTCAAACTAAAGAA 961
Qy 841 GAAGTTTCTATTAAACTACACTTAAGAGCTGTGCATAAAGAGTAACCTCACCAGAG 900
Db 962 GAGTTTCTATTAAACCACTTAAGAGCTGTGCATAAAGAGTAACCTCACCAGAA 1021
Qy 901 GACTGGATGATGTCAGCGCAGACAGTTACATTTGAATGATGCTCAACAGAGTGGAGAA 960
Db 1022 GACTGGATGATGTCAGCGCAGACAGTTACATTTGAATGATGCTCAGCCAGGTGGAGAA 1081
Qy 961 AACTGTCTGAAAAATACGCTAGAGATTGTGACACTAAGCTAGCGAGAACCAAAACAGCA 1020
Db 1082 AACCTGTCTTAAAAATACGCTAGAGATCTGTACGCTAACTAGCTAGAGAACCAAAACAGCC 1141
Qy 1021 TTTGACTTAAATTTAGAAAGCTGAACAGGAGGCAAAAGAAATCTGTTTATTTTCAATGAGCA 1200
Db 1142 TTTGACTTAAATTTAGAAAGCTGAACAGGAGGCAAAAGAAATCTGTTTATTTTCAATGAGCA 1321
Qy 1081 ACAAGAGCTTCAGAAATTTTGTCTTTTATGCTGGAATCTATGTTAAAGTTTGCCATGCT 1140
Db 1202 ACAAGAACTGTAAATTTTGTCTTTTATGCTGGAATCTATGTTAAAGTTTGCCATGCT 1261
Qy 1141 ATTTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATCTGTTTATTTTCAATGAGCA 1200
Db 1262 ATTTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATCTGTTTATTTTCAATGAGCA 1321
Qy 1201 GCCAGACAGGCAAAATCTATTATTGACAGGCAATAGCAAGAGCTTGGCAATGTTGT 1260
Db 1322 GCCAGTACAGGCAAAATCTATTATTGACAGGCAATAGCAAGAGCTTGGCAATGTTGT 1381
Qy 1261 TGCTATAATGCAGCAATGTAAACTTTTCCATTTAATGACTGTACCAACAGAACTTGAAT 1320
Db 1382 TGCTATAATGCAGCAATGTAACTTTTCCATTTAATGACTGTACCAACAGAACTTGAAT 1441
Qy 1321 TGGGTAGAAAGCTGTAACTTTTGGACAGCAAGTAACCAAGTTTAAAGCCATTTGCTCT 1380
Db 1442 TGGGTAGAAAGCTGTAACTTTTGGACAGCAAGTAACCAAGTTTAAAGCCATTTGCTCT 1501
Qy 1381 GGTCAAACTATTGCAATGATCAAAAGGAAAGGAGGCAAAAGAGTAAGTTGAAACCAACCA 1440
Db 1502 GGTCAAACTATTGCAATGATCAAAAGGAAAGGAGGCAAAAGAGTAAGTTGAAACCAACCA 1561
Qy 1441 GTCATATGACCAAAATGAGAACTTACAGTGTGAGAAATAGGCTGCGAAGAAAGACCA 1500
Db 1562 GTCATATGACCAAAATGAGAACTTACAGTGTGAGAAATAGGCTGCGAAGAGACCA 1621
Qy 1501 GAAACACCTCAACCAATCAGAGCAGAAATCTTAACTTCAATCTAACACATCTTGCCT 1560
Db 1622 GAAACACCTCAACCAATCAGAGCAGAAATCTTCAATCTAACACATCTTGCCT 1681
Qy 1561 GGTGACTTTGGTTGTAACCAAAATGAATGGCCCATGATTTGCTGTTGGTTGTAAG 1620

Db 1682 GGTGACTTTGGTTGTTGACAAAAATGAATGCGCCCATGATTTGTGCTGTTGTTGTAAG 1741
Qy 1621 AATGGTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGCAAAAGTTCTCTGATTGG 1680
Db 1742 AATGGTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGCAAAAGTTCTCTGATTGG 1801
Qy 1681 TCAGAAACCTGGCGGAGCCAAAGGTGCCAATCTCTATTAATTTACTAGGTTTCGGACCG 1740
Db 1802 TCAGAAACCTGGCGGAGCCAAAGGTGCCAATCTCTATTAATTTACTAGGTTTCGGACCG 1861
Qy 1741 TCACCAATTCAGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800
Db 1862 TCACCAATTCAGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTAACTCCACTT 1921
Qy 1801 GCATCGGATCTCGAGACCTGGCTTTAGAGCTTTGGAGCACAACCAATATACTCTGTTGCG 1860
Db 1922 GCATCGGATCTCGAGACCTGGCTTTAGAGCTTTGGAGCACAACCAATATACTCTGTTGCG 1981
Qy 1861 GGCACCTGCAGAAACCCAGAACACTGGGGAAGCTGTTCCAAAGCCTGCCAAGATGGTCAA 1920
Db 1982 GGCACCTGCAGAAACCCAGAACACTGGGGAAGCTGTTCCAAAGCCTGCCAAGATGGTCAA 2041
Qy 1921 CTGAGCCCAACTTTGGTCAGAGATCGAGGAGGATTTTCAGAGCGTGTTCGGTGGGAACCG 1980
Db 2042 CTGAGCCCAACTTTGGTCAGAGATCGAGGAGGATTTTCAGAGCGTGTTCGGTGGGAACCG 2101
Qy 1981 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019
Db 2102 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2140

RESULT 15
CQ786765
LOCUS CQ786765 5121 bp DNA linear PAT 24-MAR-2004
DEFINITION Sequence 1 from Patent WO2004018689.
ACCESSION CQ786765
VERSION CQ786765.1 GI:45721778
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Iggo, R. and Malerba, M.
TITLE Anti-neoplastic viruses
JOURNAL Patent: WO 2004018689-A 1 04-MAR-2004;
BTG INTERNATIONAL LIMITED (GB)
FEATURES
source 1..5121
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Parvovirus H1\WMV (phH1) - Parvovirus H1 with
promoter P4 and left hairpin from WMV"

ORIGIN

Query Match 87.7%; Score 1771; DB 6; Length 5121;
Best Local Similarity 92.3%; Pred. No. 0;
Matches 1864; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

Qy 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGTTTAAAGGAAAA 60
Db 282 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGTTTAAAGGAAAA 341
Qy 61 AGTAACAGAGAGTGTCTCATTTGTTTAAATAAGAAATGTTCAACTGAATGAAAA 120
Db 342 AGTAACAGAGAGTGTCTCATTTGTTTAAATAAGAAATGTTCAACTGAATGAAAA 401
Qy 121 GATATCGGATGGAATAGTTTCAAAAAAGAGCTGCGAGGAGCAGCTGAAATCTTTACAA 180
Db 402 GATATCGGATGGAATAGTTTCAAAAAAGAGCTGCGAGGAGCAGCTGAAATCTTTACAA 461
Qy 181 CGAGGAGCGAAACTACTTTGGGACCAAGGAGGACATGGAATGGGAAACCACTGGAT 240

```
Db 462 CGAGGAGCGGAACTACTTGGGACCAAGCGAGGACATGGAATGGGAAACCAACAGTGGAT 521
Qy 241 GAAATGACCAAAAAGCAAGTATTCATTTTGGTATCTTTGGTTAAATAAATGTTTATTTGAA 300
Db 522 GAAATGACCAAAAAGCAAGTATTCATTTTGGTATCTTTGGTTAAATAAATGTTTATTTGAA 581
Qy 301 GTGCTTAACCAAAAGAAATATATTTCTGGTGATGTTAAATTTGGTTGTGCAACATGAATGG 360
Db 582 GTGCTTAACCAAAAGAAATATATTTCTGGTGATGTTAAATTTGGTTGTGCAACATGAATGG 641
Qy 361 GGAAGAAGCAAGCTGGCACTGCATGTACTAAATTCGAGGAAAGGACTTTAGTCAAGCT 420
Db 642 GGAAGAAGCAAGCTGGCACTGCATGTACTAAATTCGAGGAAAGGACTTTAGTCAAGCT 701
Qy 421 CAAGGGAATTTGGTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGGTTGGTAAACAGCC 480
Db 702 CAAGGGAATTTGGTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGGTTGGTAAACAGCC 761
Qy 481 TGTAAATGTGCAACTAACCAAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAGACAAT 540
Db 762 TGTAAATGTGCAACTAACCAAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAGACAAT 821
Qy 541 GAGTGGGTTACTTACTTACTTATTAAGCATAAGCAAAACCAAAAGACTATACCAAGTGT 600
Db 822 GAGTGGGTTACTTACTTACTTATTAAGCATAAGCAAAACCAAAAGACTATACCAAGTGT 881
Qy 601 GTTCTTTTGGAAACATGATTTGCTTACTATTTTAACTAAAGAAATAAAGCACTAGT 660
Db 882 GTTCTTTTGGAAACATGATTTGCTTACTATTTTAACTAAAGAAATAAAGCACTAGT 941
Qy 661 CCACCAAGAGACGGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAACCTAACTTTTAA 720
Db 942 CCACCAAGAGACGGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAACCTAACTTTTAA 1001
Qy 721 AAAGAAGCGAGCGCCATCTAGTGAGCAAACTATACCTGATGACATGGCGCCAGAAACG 780
Db 1002 AAAGAAGCGAGCGCCATCTAGTGAGCAAACTATACCTGATGAGTGAACCCAGAAACG 1061
Qy 781 GTTGAACACACAGTAACCACTGGCGAGAAACTAAGCGCGGAGAAATTCAAACTAAAAAA 840
Db 1062 GTCGAGACACAGTGAACCACTGCACAGGAAGCTTAGCGCGGAGAAATTCAAACTAGAGAG 1121
Qy 841 GAAGTTTCTAATTAACCACTAATTAAGAGCTGGTGATTAAGAGTAACCTCACAGAG 900
Db 1122 GAGGCTCTCGATTAAACCACTCAAGAGCTGGTACATAAAAGAGTAACCTCACAGAA 1181
Qy 901 GACTGGATGATGACGACAGACAGTATCAATTGAAATGATGCTCAACCAAGTGGAGAA 960
Db 1182 GACTGGATGATGACGACGACAGTATCAATTGAAATGATGCTCAACCAAGTGGAGAA 1241
Qy 961 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTAGCCAGAACCAAAACAGCA 1020
Db 1242 AACTTGCTTAAATAACACTAGAGATCTGTACACTGACTAGCAGAGAACCAAAACAGCC 1301
Qy 1021 TTTGACTTAATTTTAGAAAGCTGAAACCAAGAAACTAACCAACTTTTCACTGCTGAC 1080
Db 1302 TTTGACTTGAATTTGGAAGAGCTGAAACCAAGAAACTAGCCAACTTTTCCATGGCTAGC 1361
Qy 1081 ACAAGAGCTGCAAGATTTTGGCTTTTCAATGGTGGAACTATGTTTAAAGTTTGGCAATGCT 1140
Db 1362 ACCAGAACCTGTGAATCTTTTGTGTAGCATGGCTGGAACTATATTAAGTCTGCCATGCC 1421
Qy 1141 ATTTGCTGTGTTTTTAAACAGACAGGAGGCAAAAGAAATACTGTTTTTATTTTATGACCA 1200
Db 1422 ATCTGTTGTGTGAATAGACAGGAGGCAAAAGGAACTACTGTGCTCTTTCACGAGCCA 1481
Qy 1201 GCAGACAGGCAAAATCTATTATGCAAGGCAATAGCAGCAAGCTTGGCAATGTTGTT 1260
Db 1482 GCCAGACAGGCAAAATCTATTATGCAAGGCAATAGCAGCAAGCTTGGTAAATGTTGTT 1541
Qy 1261 TGCTATTAATGACGCAATGTAACCTTTTCCATTTTAAATGACTGTACCAAGAACTTGATT 1320
Db 1542 TGTTACATGCTGCCAATGTGAACCTTTCCATTTTAAATGACTGTACCAAGAACTTGATT 1601
```

```
Qy 1321 TGGGTAGAAGAGCTGGTAACTTTTGGACAGCAAGTAAACCAAGCTTTTAAAGCATTTTCTCT 1380
Db 1602 TGGGTGGAAGAGCTGGTAACTTTTGGCCAGCAAGTAAACCAATTCAAAGCTATTTGTTCT 1661
Qy 1381 GGTCAAACTATTTGGCAATTTGATCAAAAAGGAAAGGAGCAGCAAAACAGATTGAACCAACCA 1440
Db 1662 GGCCAAACCATAGGCATTGATCAAAAAGGAAAGGAGCAGCAAAACAGATTGAACCAACCA 1721
Qy 1441 GTCATCATGACCAAAATGAGAACATTACAGTGGTGCAGATAGGTGCGGAAAGAACCA 1500
Db 1722 GTTATTTATGACCAACCAACGAGAACATTACCGTGGTTAGAAATAGGCTGTGAGGAAAGACCA 1781
Qy 1501 GAAACACACTCAACCAATCAGAGACAGAAATGCTTTAACTTCACTTAACATACCTTTGCCCT 1560
Db 1782 GAAACACACTCAACCAATCAGAGACAGAAATGCTTCAACATTCACCTGACAGTACACTACCT 1841
Qy 1561 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGTGCTTTGGTTGTAAG 1620
Db 1842 GGTGACTTTGGTTGGTTGATGAATGAACGAAATGGCCTCTGATCTGTGCTTTGGTTGGTGAAG 1901
Qy 1621 AATGTTTACCAATCTACCATGSCAAGCTACTGTCTTAATGGGGCAAAAGTTCTCTGATTGG 1680
Db 1902 AATGTTTACCAATCTACCATGSCCTTTTACTGTCTTAAATGGGGCAAAAGTTCTCTGATTGG 1961
Qy 1681 TCAGAAAACTGGCGGAGCCAAAAGGTGCCAACTCCTATATAATTTACTAGGTTTCGGCAGC 1740
Db 1962 TCAGAGGACTGGCGGAGCCGAAAGCTAGACACTCCTATATAATTCGCTAGGTTCAATGGCG 2021
Qy 1741 TCACCAATTCAGACACACCGAAAAAGTAGCCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800
Db 2022 TCACCAATCTGACTCCGAGAAAGTAGCCTCTCAGCCAAAACCTAGCCTCTTACTCCACTT 2081
Qy 1801 GCATCGGATCTCGAGACCTGGCTTTAGAGCCTTGGAGCAGACCAAAATACTCTCTGTTGG 1860
Db 2082 GCATCGGACCTTTGGGAGCCCTAGAGCCTTGGAGCAGACCAAAATACTCTCTGTTGG 2141
Qy 1861 GGCACCTGCAGAAACCCAGAAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920
Db 2142 GGCACCTGCAGCAAGCCAAACACTGGGGAAGCTGGTTCCACAGCCTGCCAAGGTTGCTCAA 2201
Qy 1921 CTGAGCCCAACTTTGGTTCAGAGATTCGAGGAGGATTTGAGAGCGTGTCTTGGTTCGGAACCG 1980
Db 2202 CGGAGCCCAACCTGGTTCGAGATTCGAGGCGGATTTGAGAGCTTGTCTTCAGTCAAGAAACAG 2261
Qy 1981 TTGAAGAAAGACTTCAGCGAGCCGCTGAACCTTGGACTAA 2019
Db 2262 TTGGAGAGCGACTTCAACGAGGAGCTGACCTTGGACTAA 2300
```

Search completed: January 22, 2005, 21:19:21

Job time : 8547.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2005, 23:46:57 ; Search time 950.5 Seconds

(without alignments)
11150.532 Million cell updates/sec

Title: US-10-069-056-8

Perfect score: 2019

Sequence: 1 atggctggaatgcttactc.....agccgctgaactggactaa 2019

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_23Sep04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2019	100.0	2019	5	AAD02801 Parvoviru
2	2017.4	99.9	2019	5	AAD02797 Parvoviru
3	2017.4	99.9	5149	12	Adg39767 Minute vl
4	2015.8	99.8	2019	5	AAD02805 Parvoviru
5	2015.8	99.8	2019	5	AAD02803 Parvoviru
6	2015.8	99.8	2019	5	AAD02799 Parvoviru
7	1040.4	51.5	5049	2	Aat15311 Non-atten
8	1040.4	51.5	5049	2	Aat15312 Attenuate
9	1040.4	51.5	5049	2	Aat88324 Attenuate
10	1038.8	51.5	5049	2	Aat88321 Attenuate
11	1038.8	51.5	5049	2	Aat88320 Canine pa
12	901	44.6	3524	1	Aan40252 Sequence
13	307.4	15.2	374	6	Abq95626 Tumour su
14	285.8	14.2	421	6	Abq94779 Tumour su
15	284.8	14.1	423	6	Abq94724 Tumour su
16	284.8	14.1	464	6	Abq94719 Tumour su
17	284.8	14.1	473	6	Abq94716 Tumour su
18	284.8	14.1	486	6	Abq94707 Tumour su
19	284.8	14.1	491	6	Abq94777 Tumour su
20	284.2	14.1	420	6	Abq94787 Tumour su
21	283.8	14.1	343	6	Abq94737 Tumour su

22	283.8	14.1	343	6	ABQ94739	Tumour su
23	283.8	14.1	403	6	ABQ94782	Tumour su
24	283.2	14.0	497	6	ABQ94718	Tumour su
25	280.4	13.9	340	6	ABQ94732	Tumour su
26	278.2	13.8	342	6	ABQ94765	Tumour su
27	273.8	13.6	420	6	ABQ94781	Tumour su
28	269.6	13.4	324	6	ABQ94740	Tumour su
29	263.4	13.0	339	6	ABQ95625	Tumour su
30	261.6	13.0	424	6	ABQ94778	Tumour su
31	163.8	8.1	468	6	ABQ94873	Tumour su
32	162.2	8.0	451	6	ABQ94868	Tumour su
33	162.2	8.0	457	6	ABQ94867	Tumour su
34	162.2	8.0	465	6	ABQ94874	Tumour su
35	162.2	8.0	469	6	ABQ94869	Tumour su
36	162.2	8.0	472	6	ABQ94880	Tumour su
37	162.2	8.0	515	6	ABQ94793	Tumour su
38	162.2	8.0	516	6	ABQ94807	Tumour su
39	162.2	8.0	516	6	ABQ94794	Tumour su
40	162.2	8.0	516	6	ABQ94800	Tumour su
41	162.2	8.0	517	6	ABQ94811	Tumour su
42	162.2	8.0	530	6	ABQ94791	Tumour su
43	162.2	8.0	530	6	ABQ94799	Tumour su
44	161	8.0	516	6	ABQ94796	Tumour su
45	160.6	8.0	530	6	ABQ94805	Tumour su

ALIGNMENTS

RESULT 1

AAD02801

ID AAD02801 standard; DNA; 2019 BP.

XX AC

XX AC

DT 06-AUG-2003 (revised)

DT 31-MAY-2001 (first entry)

XX XX

DE Parvovirus non-structure protein 1 (NS1) variant (T363A) DNA.

KW NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
tumoural disease; gene therapy; mutant; mutein; variant; ds.

XX Parvovirus.

OS Synthetic.

XX XX

FH Key Location/Qualifiers

FT CDS 1..2019

FT /tag= a

FT mutation /product= "NS1 variant (T363A) protein"

FT /replace(1187, A)

FT /*tag= b

XX EPI077260-A1.

XX 21-FEB-2001.

XX 13-AUG-1999; 99EP-00115161.

XX 13-AUG-1999; 99EP-00115161.

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX Nueesch J, Rommelaere J;

XX WFI; 2001-212717/22.

XX P-PSDB; AAY72706.

XX Novel parvovirus non-structure protein variant, useful for treating
tumoral diseases, has a shifted equilibrium between DNA replication and
transcription activities, and cytotoxic activity.

XX Claim 7; Page 16-19; 41pp; English.

QY	1981	TTGAAAGAAAGACTTCAGCGAGCGCTGAACCTTGGA	2019
Db	1981	TTGAAGAAAGACTTCAGCGAGCGCTGAACCTTGGA	2019
RESULT 2			
AD02797	AAD02797 standard; DNA; 2019 BP.		
ID	AAD02797		
AC	06-AUG-2003 (revised)		
XX	31-MAY-2001 (first entry)		
DT	Parvovirus non-structure protein 1 (NS1) wild-type DNA.		
DE	NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;		
XX	tumoural disease; gene therapy; ds.		
KW	Parvovirus.		
OS	Key		
XX	Location/Qualifiers		
XX	1. 2019		
FT	/*tag= a		
FT	/product= "Parvovirus NS1 protein"		
XX	EP1077260-A1.		
PN	21-FEB-2001.		
XX	13-AUG-1999; 99EP-00115161.		
XX	13-AUG-1999; 99EP-00115161.		
XX	(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.		
PA	Nuesch J, Rommelaere J;		
XX	WPI; 2001-212717/22.		
PI	P-PSDB; AAY72702.		
XX	Novel parvovirus non-structure protein variant, useful for treating		
PT	tumoral diseases, has a shifted equilibrium between DNA replication and		
PT	transcription activities, and cytotoxic activity.		
XX	Disclosure; Fig 1; 41pp; English.		
PS	The present sequence is a wild type DNA encoding parvovirus non-		
XX	structure protein 1 (NS1). The present invention relates to the variants		
CC	of the parvovirus non-structure protein (NS1) having a shifted		
CC	equilibrium between the DNA replication and transcription activities, and		
CC	the cytotoxicity activity. These variants are useful as toxins for		
CC	treating tumoural diseases. The variant DNAs are useful as vectors for		
XX	gene therapy. (Updated on 06-AUG-2003 to correct OS field.)		
XX	Sequence 2019 BP; 698 A; 389 C; 452 G; 480 T; 0 U; 0 Other;		
SQ			
Query Match 99.9%; Score 2017.4; DB 5; Length 2019;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 2018; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	1	ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA	60
Db	1	ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA	60
QY	61	AGTAACCAAGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAA	120
Db	61	AGTAACCAAGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAA	120
QY	121	GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGCAGCTGAAATCTTTTACAA	180
Db	121	GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGCAGCTGAAATCTTTTACAA	180
QY	181	CGAGGAGCGGAAACTTACTTTGGGACCAAGCGAGGACATGGAATGGGAAACCAAGTGGAT	240
Db	181	CGAGGAGCGGAAACTTACTTTGGGACCAAGCGAGGACATGGAATGGGAAACCAAGTGGAT	240
QY	241	GAAATGACCAAAAGCAAGTATTCTTTTGGTATTCTTTGTTTAAAAAATGTTTATTGAA	300
Db	241	GAAATGACCAAAAGCAAGTATTCTTTTGGTATTCTTTGTTTAAAAAATGTTTATTGAA	300
QY	301	GTGCTTAAACAAAGAAATATTTCTGCTGATGTTAAATTTGTTTGTCAACATGAATGG	360
Db	301	GTGCTTAAACAAAGAAATATTTCTGCTGATGTTAAATTTGTTTGTCAACATGAATGG	360
QY	361	GGAAAAGCAAGGCTGGCAGCTGCCATGTACTTAATTTGGAGGAAAGGACTTTTAGCTCAAGCT	420
Db	361	GGAAAAGCAAGGCTGGCAGCTGCCATGTACTTAATTTGGAGGAAAGGACTTTTAGCTCAAGCT	420
QY	421	CAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTGAGCAGATGTTGGTAAACAGCC	480
Db	421	CAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTGAGCAGATGTTGGTAAACAGCC	480
QY	481	TGTAATGTGCAACTAACACAGCTGAAGAAATTTAACTAAGAGAAATAGCAGAGCAAT	540
Db	481	TGTAATGTGCAACTAACACAGCTGAAGAAATTTAACTAAGAGAAATAGCAGAGCAAT	540
QY	541	GAGTGGTTTACTCTACTTATTAAGCATAAAGCAACCAAAAGAACTATACCAAGTGT	600
Db	541	GAGTGGTTTACTCTACTTATTAAGCATAAAGCAACCAAAAGAACTATACCAAGTGT	600
QY	601	GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAGAAATAAGCACTAGT	660
Db	601	GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAGAAATAAGCACTAGT	660
QY	661	CCACCAAGAGAGCGAGGCTATTTCTTAGCAGTACTCTGGCTGGAATACTTCTTTTA	720
Db	661	CCACCAAGAGAGCGAGGCTATTTCTTAGCAGTACTCTGGCTGGAATACTTCTTTTA	720
QY	721	AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGGGCCAGAAACG	780
Db	721	AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGGGCCAGAAACG	780
QY	781	GTTGAAACCAAGTAACTACACTTAAAGAGCTGGTGCATATAAAGAGTAACTTCAACAGAG	840
Db	781	GTTGAAACCAAGTAACTACACTTAAAGAGCTGGTGCATATAAAGAGTAACTTCAACAGAG	840
QY	841	GAAATTTCTATTAAAACTACACTTAAAGAGCTGGTGCATATAAAGAGTAACTTCAACAGAG	900
Db	841	GAAATTTCTATTAAAACTACACTTAAAGAGCTGGTGCATATAAAGAGTAACTTCAACAGAG	900
QY	901	GACTGATGATGATGAGCAGACAGTTTACATTGAAATGATGGCTCAACAGGTGGAGAA	960
Db	901	GACTGATGATGATGAGCAGACAGTTTACATTGAAATGATGGCTCAACAGGTGGAGAA	960
QY	961	AACCTGCTGAAAAATACGCTAGAGATTTGTACACTTAACTCTAGCCAGAACCAACAGCA	1020
Db	961	AACCTGCTGAAAAATACGCTAGAGATTTGTACACTTAACTCTAGCCAGAACCAACAGCA	1020
QY	1021	TTTGACTTAAATTTAGAAAAAGCTGAAACCAAGCAAACTAACCACTTTTCACTGCCTGAC	1080
Db	1021	TTTGACTTAAATTTAGAAAAAGCTGAAACCAAGCAAACTAACCACTTTTCACTGCCTGAC	1080
QY	1081	ACAAGAGCTGCAAGATTTTGTCTTTTGTGGTGGAACTATGTTAAAGTTTGCATGCT	1140
Db	1081	ACAAGAGCTGCAAGATTTTGTCTTTTGTGGTGGAACTATGTTAAAGTTTGCATGCT	1140
QY	1141	ATTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATACTGTTTATTTATGAGCA	1200
Db	1141	ATTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATACTGTTTATTTATGAGCA	1200
QY	1201	GCCAGCAGAGGCAAACTCTATTATTGCAAGCCATAGCAAGCAGTTGGCAATGTTGGT	1260
Db	1201	GCCAGCAGAGGCAAACTCTATTATTGCAAGCCATAGCAAGCAGTTGGCAATGTTGGT	1260

```
QY 1261 TGCTATTAATGAGCAATGAACTTTCCATTTAATGACTGTACCAACAAGAACTTGATT 1320
Db 1261 TGCTATTAATGAGCAATGAACTTTCCATTTAATGACTGTACCAACAAGAACTTGATT 1320
QY 1321 TGGGTAGAAGAGCTGTTGCTTGGACAGCAAGTAAACAGTTTAAAGCCATTGCTCT 1380
Db 1321 TGGGTAGAAGAGCTGTTGCTTGGACAGCAAGTAAACAGTTTAAAGCCATTGCTCT 1380
QY 1381 GGTCAAACTATTCCGATTGATCAAAAAGGAAAGGAGCAGCAAAACAGATTGAAACCAACCA 1440
Db 1381 GGTCAAACTATTCCGATTGATCAAAAAGGAAAGGAGCAGCAAAACAGATTGAAACCAACCA 1440
QY 1441 GTCATCATGACCAAAATGAAACATTACAGTGTGAGTAAGCTGCGAAGAAAGACCA 1500
Db 1441 GTCATCATGACCAAAATGAAACATTACAGTGTGAGTAAGCTGCGAAGAAAGACCA 1500
QY 1501 GAACACACTCAACCAATCAGACAGCAAGATGCTTAAACATTCATCTAACACATACCTTGCTT 1560
Db 1501 GAACACACTCAACCAATCAGACAGCAAGATGCTTAAACATTCATCTAACACATACCTTGCTT 1560
QY 1561 GGTGACTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGCTTGGTTGGTAAAG 1620
Db 1561 GGTGACTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGCTTGGTTGGTAAAG 1620
QY 1621 AATGGTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTGATTGG 1680
Db 1621 AATGGTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTGATTGG 1680
QY 1681 TCAGAAACTGGGGAGGCAAAAGTCCCACTTCTTAAATTTACTAGTTGGGACCG 1740
Db 1681 TCAGAAACTGGGGAGGCAAAAGTCCCACTTCTTAAATTTACTAGTTGGGACCG 1740
QY 1741 TCACCATTCAGACACCGAAAGTACGCCCTCTCAGCCAGAACTATGCACTCACTTCCACTT 1800
Db 1741 TCACCATTCAGACACCGAAAGTACGCCCTCTCAGCCAGAACTATGCACTCACTTCCACTT 1800
QY 1801 GCATCGGATCTGAGGACCTGGCTTTAGAGCTTTGAGGACACCAAAATCTCTGTTGG 1860
Db 1801 GCATCGGATCTGAGGACCTGGCTTTAGAGCTTTGAGGACACCAAAATCTCTGTTGG 1860
QY 1861 GGCACCTGCAGAAACCCAGACACTGGGAGCTGTTTCCAAAGCCTGCCAAGATGCTCAA 1920
Db 1861 GGCACCTGCAGAAACCCAGACACTGGGAGCTGTTTCCAAAGCCTGCCAAGATGCTCAA 1920
QY 1921 CTGAGGCCCACTTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTGGTGGGAAACCG 1980
Db 1921 CTGAGGCCCACTTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTGGTGGGAAACCG 1980
QY 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2019
Db 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2019
```

RESULT 3

```
ADG39767
ID ADG39767 standard; DNA; 5149 BP.
XX
AC ADG39767;
XX
DT 11-MAR-2004 (first entry)
XX
DE Minute virus from mouse genomic DNA sequence SEQ ID NO:29.
XX
KW parvovirus; rep; cap; DNA binding domain; capsid interacting domain;
KW recombinant hybrid parvovirus particle;
KW recombinant adeno-associated virus; rAAV; AAV; vaccine; gene therapy;
KW gene; ds.
XX
OS Mice minute virus.
XX
PN WO2003104392-A2.
XX
PD 18-DEC-2003.
```

```
XX 02-DEC-2002; 2002WO-US038423.
PF
XX
PR 18-DEC-2001; 2001US-0341919P.
XX
PA (UYN-) UNIV NORTH CAROLINA.
XX
PI Samulski RJ, Rabinowitz JE;
XX
PI WPI; 2004-062324/06.
XX
PT New polynucleotides comprising parvovirus rep coding sequences and
PT parvovirus cap coding sequences, useful in producing higher stocks of
PT hybrid parvovirus vectors for delivering therapeutic nucleic acids to a
PT subject.
XX
PS Disclosure; SEQ ID NO 29; 115pp; English.
CC
CC The present invention describes a polynucleotide (I), comprising
CC parvovirus rep coding sequences and parvovirus cap coding sequences. The
CC rep coding sequences encodes a DNA binding domain from a first
CC parvovirus, and a capsid interacting domain from a parvovirus different
CC from the first parvovirus. The cap coding sequence comprises sequences
CC from the different parvovirus. Also described: (1) a vector comprising
CC (1); (2) a cell comprising (1), or parvovirus rep coding sequences and
CC parvovirus cap coding sequences, where the rep coding sequences encode a
CC DNA binding domain from a first parvovirus and a capsid interacting
CC domain from a parvovirus different from the first parvovirus; the cap
CC coding sequences comprise sequences from the different parvovirus, and
CC the rep coding sequences are stably integrated into the genome of the
CC cell; and (3) producing a recombinant hybrid parvovirus particle or adeno
CC -associated virus (rAAV) particle. (I) can be used in vaccines, and in
CC gene therapy. The polynucleotide (I) can be used in producing higher
CC stocks of hybrid parvoviruses or parvovirus vectors, which may be used in
CC the delivery of nucleic acids having biological effect to treat or
CC ameliorate the symptoms associated with any disorder related to gene
CC expression. The polynucleotide may be used to produce a parvovirus vector
CC to express an immunogenic polypeptide in a subject, e.g. for vaccination.
CC The parvovirus vector may also be used to provide an antisense nucleic
CC acid to a cell in vitro or in vivo, or in diagnostic and screening
CC methods. The present sequence is used in the exemplification of the
CC present invention.
XX
SQ Sequence 5149 BP; 1718 A; 1045 C; 1124 G; 1262 T; 0 U; 0 Other;
Query Match 99.9%; Score 2017.4; DB 12; Length 5149;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2018; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 60
Db 261 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 320
QY 61 AGTAACCGAGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAA 120
Db 321 AGTAACCGAGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAA 380
QY 121 GATATCGGATGAATAGTTTACAAAAAGAGCTGCGAGGAGGAGCTGAAATCTTTACAA 180
Db 381 GATATCGGATGAATAGTTTACAAAAAGAGCTGCGAGGAGGAGCTGAAATCTTTACAA 440
QY 181 CGAGGAGCGGAACTACTTGGGACCAAGCGAGGACATGGATGGGAAACCAACAGTGGAT 240
Db 441 CGAGGAGCGGAACTACTTGGGACCAAGCGAGGACATGGATGGGAAACCAACAGTGGAT 500
QY 241 GAAATGACCAAAACAGTATTCTATTTGATTTGTTTGGTTAAAAATGTTTATTGAA 300
Db 501 GAAATGACCAAAACAGTATTCTATTTGATTTGTTTGGTTAAAAATGTTTATTGAA 560
QY 301 GTGCTTAAACAAAGATATATTTCTCGTGTATGTTAATTTGTTTGTGCAACATGAATGG 360
Db 561 GTGCTTAAACAAAGATATATTTCTCGTGTATGTTAATTTGTTTGTGCAACATGAATGG 620
```

361 GGAAGAAGCCAGGCTGGCACTGCCATGTACTAATTTGGAGGAAGGACTTTTGTCAAGCT 420
Db |||||
621 GGAAGAAGCCAGGCTGGCACTGCCATGTACTAATTTGGAGGAAGGACTTTTGTCAAGCT 680
Qy |||||
421 CAAAGGAAATGGTGGAGAAAGGCACTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCC 480
Db |||||
681 CAAAGGAAATGGTGGAGAAAGGCACTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCC 740
Qy |||||
481 TGTAAATGTGCAACTAACACAGCTGAAAGAAATTAATTAAGAGAAATAGCAGAGAAAT 540
Db |||||
741 TGTAAATGTGCAACTAACACAGCTGAAAGAAATTAATTAAGAGAAATAGCAGAGAAAT 800
Qy |||||
541 GAGTGGTGTACTTACTTATATAGCATAAGCAACCAAAAGAGACTATACCAGTGT 600
Db |||||
801 GAGTGGTGTACTTACTTATATAGCATAAGCAACCAAAAGAGACTATACCAGTGT 860
Qy |||||
601 GTTCTTTTGGAAACATGATGTCTTACTATTTTAACTAAAGAGAAATAGCACTAGT 660
Db |||||
861 GTTCTTTTGGAAACATGATGTCTTACTATTTTAACTAAAGAGAAATAGCACTAGT 920
Qy |||||
661 CCACCAAGAGACGGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTTAACCTTTTAA 720
Db |||||
921 CCACCAAGAGACGGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTTAACCTTTTAA 980
Qy |||||
721 AAGAAGGCGAGCGCCATCTAGTAGCAAACTATACACTGATGACATGCGGCCAGAAACG 780
Db |||||
981 AAGAAGGCGAGCGCCATCTAGTAGCAAACTATACACTGATGACATGCGGCCAGAAACG 1040
Qy |||||
781 GTTGAACACACAGTAAACCACTGGCGAGGAACTTAAGCGGGCAGAAATCAAACTAAAAAA 840
Db |||||
1041 GTTGAACACACAGTAAACCACTGGCGAGGAACTTAAGCGGGCAGAAATCAAACTAAAAAA 1100
Qy |||||
841 GAAATTTCTATTAAAACTACACTTAAGAGCTGGTGCATAAAAAGAGTAACCTCACAGAG 900
Db |||||
1101 GAAATTTCTATTAAAACTACACTTAAGAGCTGGTGCATAAAAAGAGTAACCTCACAGAG 1160
Qy |||||
901 GACTGATGATGATGACGAGCAGACAGTACATTGAAATGATGCTCAACCGAGTGGAGAA 960
Db |||||
1161 GACTGATGATGATGACGAGCAGACAGTACATTGAAATGATGCTCAACCGAGTGGAGAA 1220
Qy |||||
961 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTTAGCCAGAAACCAAAACAGCA 1020
Db |||||
1221 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTTAGCCAGAAACCAAAACAGCA 1280
Qy |||||
1281 TTTGACTTAAATTTAGAAAAAGCTGAAACCCAGCAAACTAACCACTTTTCACTGCTGAC 1080
Db |||||
1081 ACAAGAGCTGCAAGATTTTGGCTTTTTCATGGCTGGAATATGTTAAAGTTTGGCATGCT 1140
Qy |||||
1341 ACAAGAGCTGCAAGATTTTGGCTTTTTCATGGCTGGAATATGTTAAAGTTTGGCATGCT 1400
Db |||||
1141 ATTTGCTGTGTTTAAACAGACAGAGGCAAAAGAAATACTGTTTATTTATGACCA 1200
Qy |||||
1401 ATTTGCTGTGTTTAAACAGACAGAGGCAAAAGAAATACTGTTTATTTATGACCA 1460
Db |||||
1201 GCCAGCAGCGCAAACTATTATTGCAAGCCATAGCAACAGCAGTTGGCAATGTTGGT 1260
Qy |||||
1461 GCCAGCAGCGCAAACTATTATTGCAAGCCATAGCAACAGCAGTTGGCAATGTTGGT 1520
Db |||||
1261 TGCTATAATGACGCAATGTAACCTTTCCATTTAATGACTGTACCAAGAACTTTGATT 1320
Qy |||||
1521 TGCTATAATGACGCAATGTAACCTTTCCATTTAATGACTGTACCAAGAACTTTGATT 1580
Db |||||
1321 TGGGTAGAAGAGCTGTACTTTTGGACAGCAAGTAAACAGTTTAAAGCATTTGCTCT 1380
Qy |||||
1581 TGGGTAGAAGAGCTGTACTTTTGGACAGCAAGTAAACAGTTTAAAGCATTTGCTCT 1640
Db |||||
1381 GGTCAAACTATTGCTATTGATCAAAAGAGAAAGGAGCAAAACAGATTGAACCAACACCA 1440
Qy |||||
1641 GGTCAAACTATTGCTATTGATCAAAAGAGAAAGGAGCAAAACAGATTGAACCAACACCA 1700
Db |||||
1441 GTCATCATGACCAAAATAGAAACATTACAGTGGTCAGAAATAGGCTGCGGAAGAAAGACCA 1500
Qy |||||

1701 GTCATCATGACCAAAATAGAAACATTACAGTGGTCAGATAGGCTGCGAAGAAAGACCA 1760
Qy |||||
1501 GAAACACACTCAACCAATCAGAGACAGAATGCTTTAATCATTCATCTTAACACATACCTTGCCT 1560
Db |||||
1761 GAAACACACTCAACCAATCAGAGACAGAATGCTTTAATCATTCATCTTAACACATACCTTGCCT 1820
Qy |||||
1561 GGTGACTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGTCTTGGTTGGTAAAG 1620
Db |||||
1821 GGTGACTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGTCTTGGTTGGTAAAG 1880
Qy |||||
1621 AATGGTTACCAATCTACCATGGCAAGCTACTGTCTTAATGGGGCAAAAGTTCTCTGATTGG 1680
Db |||||
1881 AATGGTTACCAATCTACCATGGCAAGCTACTGTCTTAATGGGGCAAAAGTTCTCTGATTGG 1940
Qy |||||
1681 TCAGAAAACTGGCGGAGCCAAAAGGTGCCAACTCTCTATAAATTTACTAGTTGCGGCACGC 1740
Db |||||
1941 TCAGAAAACTGGCGGAGCCAAAAGGTGCCAACTCTCTATAAATTTACTAGTTTGGGCACGC 2000
Qy |||||
1741 TCACCAATTCAGACACCCGAAAGTAGCGCTCTCAGCCAGAACTATGSCACTTAATCCACTT 1800
Db |||||
2001 TCACCAATTCAGACACCCGAAAGTAGCGCTCTCAGCCAGAACTATGSCACTTAATCCACTT 2060
Qy |||||
1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACACCAAAATACTCTCTGTGGC 1860
Db |||||
2061 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACACCAAAATACTCTCTGTGGC 2120
Qy |||||
1861 GGCACCTGCAGAAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGTCAT 1920
Db |||||
2121 GGCACCTGCAGAAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGTCAT 2180
Qy |||||
1921 CTGAGCCCAACTTGTTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGTTGGGAACCG 1980
Db |||||
2181 CTGAGCCCAACTTGTTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGTTGGGAACCG 2240
Qy |||||
1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTTGGACTAA 2019
Db |||||
2241 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTTGGACTAA 2279

RESULT 4
AAD02805
ID AAD02805 standard; DNA; 2019 BP.
XX
AC AAD02805;
XX
DT 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
DE Parvovirus non-structure protein 1 (NS1) variant (T463A) DNA.
KW NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
tumoural disease; gene therapy; mutant; mutein; variant; ds.
XX
OS Parvovirus.
OS Synthetic.
XX
FH Key
CDS 1. .2019
FT /tag= a
FT /product= "NS1 variant (T463A) protein"
FT replace(1387, A)
FT /tag= b
XX
PN EP1077260-A1.
XX
PD 21-FEB-2001.
XX
PF 13-AUG-1999; 99EP-00115161.
XX
PR 13-AUG-1999; 99EP-00115161.
XX
PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

[illegible]

QY 1801 GCATCGCATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACCAATATCTCTGTGGC 1860
Db 1801 GCATCGCATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACCAATATCTCTGTGGC 1860
QY 1861 GGCACTCGAGAAACCCAGAACACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920
Db 1861 GGCACTCGAGAAACCCAGAACACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920
QY 1921 CTGAGCCCAACTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGTTCGGTGGGGAACCG 1980
Db 1921 CTGAGCCCAACTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGTTCGGTGGGGAACCG 1980
QY 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACCTGGACTAA 2019
Db 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACCTGGACTAA 2019

RESULT 5

AA02803
ID AAD02803 standard; DNA; 2019 BP.
XX AC
AC AC
XX XX
DT 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX XX
DE Parvovirus non-structure protein 1 (NS1) variant (T394A) DNA.
XX XX
KW NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoral disease; gene therapy; mutant; muten; variant; ds.
XX XX

Parvovirus.

OS Synthetic.
XX XX
FH Key Location/Qualifiers
CDS 1..2019
FT /*tag= a
FT /product= "NS1 variant (T394A) protein"
FT replace(1180, A)
FT /*tag= b
XX XX
PN EP1077260-A1.
XX XX
PD 21-FEB-2001.
XX XX
PF 13-AUG-1999; 99EP-00115161.
XX XX
PR 13-AUG-1999; 99EP-00115161.
XX XX
PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX XX

P1 Nueesch J, Roumelaere J;

PI WPI; 2001-212717/22.

DR P-PSDB; AAY72708.

XX XX
PT Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX XX
PS Claim 7; Page 22-24; 41pp; English.
XX XX
CC The present sequence is a DNA encoding parvovirus non-structure protein 1
CC (NS1) variant (T394A). The invention relates to the variants of the
CC parvovirus non-structure protein (NS1) having a shifted equilibrium
CC between the DNA replication and transcription activities, and the
CC cytotoxicity activity. These variants are useful as toxins for treating
CC tumoral diseases. The variant DNAs are useful as vectors for gene
CC therapy. (Updated on 06-AUG-2003 to correct OS field.)
XX XX

QY Sequence 2019 BP; 697 A; 389 C; 453 G; 480 T; 0 U; 0 Other;

Query Match 99.8%; Score 2015.8; DB 5; Length 2019;

Best Local Similarity 99.9%; Pred. No. 0;
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGCGCTGGAAATGCTTACTCTGATGAAGTTTTCGGAGCAACCACTGGTTTAAAGGAAAA 60
Db 1 ATGCGCTGGAAATGCTTACTCTGATGAAGTTTTCGGAGCAACCACTGGTTTAAAGGAAAA 60
QY 61 AGTAACAGAGAGTGTCTTCAATTTGTTTTTAAAAATGAAATGTTCAACTGAATGGAAAA 120
Db 61 AGTAACAGAGAGTGTCTTCAATTTGTTTTTAAAAATGAAATGTTCAACTGAATGGAAAA 120
QY 121 GATATCGGATGGAAATAGTTTACAAAAAGAGCTCGAGAGGACGAGCTGAAATCTTTACAA 180
Db 121 GATATCGGATGGAAATAGTTTACAAAAAGAGCTCGAGAGGACGAGCTGAAATCTTTACAA 180
QY 181 CGAGGAGCGGAAACTACTTTGGGACCAAGCGAGGACATGGAATGGGAAAAACCAAGTGGAT 240
Db 181 CGAGGAGCGGAAACTACTTTGGGACCAAGCGAGGACATGGAATGGGAAAAACCAAGTGGAT 240
QY 241 GAAATGACCAAAAAAGCAAGTATTTTATTTGTTTAAATGTTTAAATGTTTATTTGAA 300
Db 241 GAAATGACCAAAAAAGCAAGTATTTTATTTGTTTAAATGTTTAAATGTTTATTTGAA 300
QY 301 GTGCTTAAACACAAAGAAATATTTCTCTGGTGTATTTAAATGTTTGTGCAACATGAATGG 360
Db 301 GTGCTTAAACACAAAGAAATATTTCTCTGGTGTATTTAAATGTTTGTGCAACATGAATGG 360
QY 361 GGAAGAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGACTTTTGTCAAGCT 420
Db 361 GGAAGAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGACTTTTGTCAAGCT 420
QY 421 CAAGGAAATGTTGGAGAAAGGCAACTAAATGTTTACTGGAGCAGATGGTTGTTAAACAGCC 480
Db 421 CAAGGAAATGTTGGAGAAAGGCAACTAAATGTTTACTGGAGCAGATGGTTGTTAAACAGCC 480
QY 481 TGTATGTGCAACTAACACAGCTGAAGAAATTTAAATTAAGCAATATAGCAATAGCAGAACAT 540
Db 481 TGTATGTGCAACTAACACAGCTGAAGAAATTTAAATTAAGCAATATAGCAGAACAT 540
QY 541 GAGTGGTTTACTCTTACTTATTAAGCATTAAGCAACCAAAAAAGACTATATCAAGTGT 600
Db 541 GAGTGGTTTACTCTTACTTATTAAGCATTAAGCAACCAAAAAAGACTATATCAAGTGT 600
QY 601 GTTCTTTTGGAAACATGATTTCTTACTTATTTTAACTTAAAGGAAATTAAGCACTAGT 660
Db 601 GTTCTTTTGGAAACATGATTTCTTACTTATTTTAACTTAAAGGAAATTAAGCACTAGT 660
QY 661 CCACCAAGAGACGGAGGCTATTTTCTTAGCAGTACTCTGGCTGGGAAACTAATCTTTTAA 720
Db 661 CCACCAAGAGACGGAGGCTATTTTCTTAGCAGTACTCTGGCTGGGAAACTAATCTTTTAA 720
QY 721 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGGGCCAGAAACG 780
Db 721 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGGGCCAGAAACG 780
QY 781 GTTGAACACACAGTAAACCACTGCGCAGGAAACTAAGCGCGCAGAAATTCAAACTTAAAAA 840
Db 781 GTTGAACACACAGTAAACCACTGCGCAGGAAACTAAGCGCGCAGAAATTCAAACTTAAAAA 840
QY 841 GAAAGTTTCTTAAACTACACTTAAAGAGCTGGTGCATAAAAAGAGTAACTCACCAGAG 900
Db 841 GAAAGTTTCTTAAACTACACTTAAAGAGCTGGTGCATAAAAAGAGTAACTCACCAGAG 900
QY 901 GACTGATGATCATGACGAGCAGACAGTTCATTTGAATGATGCTCAACACAGTGGAGAA 960
Db 901 GACTGATGATCATGACGAGCAGACAGTTCATTTGAATGATGCTCAACACAGTGGAGAA 960
QY 961 AACCTGCTGAAAAATACGCTAGAGATTTGTACAACTTAACCTCTAGCCAGAAACCAACAGCA 1020
Db 961 AACCTGCTGAAAAATACGCTAGAGATTTGTACAACTTAACCTCTAGCCAGAAACCAACAGCA 1020
QY 1021 TTTGACTTAAATTTAGAAAAAGCTGAAACCAAGCAACCTAACTTTTCACTGCTCTGAC 1080
Db 1021 TTTGACTTAAATTTAGAAAAAGCTGAAACCAAGCAACCTAACTTTTCACTGCTCTGAC 1080

QY 301 GTGCTTAACAAGAAATATTTCTGGTGATGTTAAATGGTTTGTGCAACATGAATGG 360
Db 301 GTGCTTAACAAGAAATATTTCTGGTGATGTTAAATGGTTTGTGCAACATGAATGG 360
QY 361 GGAAGAGACCAAGGCTGGCACTGCCATGTACTAATTTGGAGAAAGCACTTTAGTCAAGCT 420
Db 361 GGAAGAGACCAAGGCTGGCACTGCCATGTACTAATTTGGAGAAAGCACTTTAGTCAAGCT 420
QY 421 CAAGGGAATTTGGTGGAGAGGCAAACTAAATGTTTACTGGAGCAGATGGTTGTTAAAGCC 480
Db 421 CAAGGGAATTTGGTGGAGAGGCAAACTAAATGTTTACTGGAGCAGATGGTTGTTAAAGCC 480
QY 481 TGTAAATGTGCAACTACACACAGCTGAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT 540
Db 481 TGTAAATGTGCAACTACACACAGCTGAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT 540
QY 541 GAGTGGGTACTCTACTTACTTATAGCATAAGCAAAACCAAAAGCACTATACCAAGTGT 600
Db 541 GAGTGGGTACTCTACTTACTTATAGCATAAGCAAAACCAAAAGCACTATACCAAGTGT 600
QY 601 GTTCTTTTGGAAACATGATGTCTTACTATTTTAACTTAAAGAAATAAAGCACTAGT 660
Db 601 GTTCTTTTGGAAACATGATGTCTTACTATTTTAACTTAAAGAAATAAAGCACTAGT 660
QY 661 CCACCAAGAGACGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAACTAATCTTTTA 720
Db 661 CCACCAAGAGACGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAACTAATCTTTTA 720
QY 721 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG 780
Db 721 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG 780
QY 781 GTTGAACCAACAGTAACCACTGGCGAGGAACTAAGCGCGGAGAAATTCAACTAAAAAA 840
Db 781 GTTGAACCAACAGTAACCACTGGCGAGGAACTAAGCGCGGAGAAATTCAACTAAAAAA 840
QY 841 GAAGTTTCTAATTAACACTACATTAAGAGCTGGTGCAATAAAGATRACTCACCAGAG 900
Db 841 GAAGTTTCTAATTAACACTACATTAAGAGCTGGTGCAATAAAGATRACTCACCAGAG 900
QY 901 GACTGGATGATGACGACGACAGTTTACATTAAGATGATGCTCAACACAGGTGGAGAA 960
Db 901 GACTGGATGATGACGACGACAGTTTACATTAAGATGATGCTCAACACAGGTGGAGAA 960
QY 961 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAAACCAAAACAGCA 1020
Db 961 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAAACCAAAACAGCA 1020
QY 1021 TTTGACTTAATTTAGAAAAAGCTGAAACAGCAAACTAACCACTTTTCACTGCCTGAC 1080
Db 1021 TTTGACTTAATTTAGAAAAAGCTGAAACAGCAAACTAACCACTTTTCACTGCCTGAC 1080
QY 1081 ACAAGACCTGCAAGATTTTGTCTTTTCATGCTGGNACTATGTTAAAGTTGCCATGCT 1140
Db 1081 ACAAGACCTGCAAGATTTTGTCTTTTCATGCTGGNACTATGTTAAAGTTGCCATGCT 1140
QY 1141 ATTTGCTGTGTTTAAACAGACAGGAGCAAAAGAAATACTGTTTATTTTATGAGACCA 1200
Db 1141 ATTTGCTGTGTTTAAACAGACAGGAGCAAAAGAAATACTGTTTATTTTATGAGACCA 1200
QY 1201 GCCAGCACAGGCAAACTTATTTGCAAAAGCCATAGCACAAGCAGTTGGCAATGTTGT 1260
Db 1201 GCCAGCACAGGCAAACTTATTTGCAAAAGCCATAGCACAAGCAGTTGGCAATGTTGT 1260
QY 1261 TGCTATAATCGAGCAATGTAACCTTTTCACTTTAATGACTGTACCAACAGAACTTGATT 1320
Db 1261 TGCTATAATCGAGCAATGTAACCTTTTCACTTTAATGACTGTACCAACAGAACTTGATT 1320
QY 1321 TGGTGAAGAGCTGGTAACTTTGGACAGCAGTAACCAAGTTTAAAGCACTTTGCTCT 1380
Db 1321 TGGTGAAGAGCTGGTAACTTTGGACAGCAGTAACCAAGTTTAAAGCACTTTGCTCT 1380
QY 1381 GGTCAAACTATTGGCTATGATCAAAAAGGAAAGGCGAGCAAAACAGATTTGAACCAACCA 1440

Db 1381 GGTCAAACTATTGGCTATTGCTATTAATGCTTAAATGGTTTGTGCAACATGAATGG 1440
QY 1441 GTCATCATGACCAAAATCAGAACTATTACAGTGGTTCAGAAATAGGCTGCGAAGAAAGACCA 1500
Db 1441 GTCATCATGACCAAAATCAGAACTATTACAGTGGTTCAGAAATAGGCTGCGAAGAAAGACCA 1500
QY 1501 GAACACACTCAACCAATCAGAGACAGAAATGCTTAACTTCACTTAACATACATCTTGCCT 1560
Db 1501 GAACACACTCAACCAATCAGAGACAGAAATGCTTAACTTCACTTAACATACATCTTGCCT 1560
QY 1561 GGTGACTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGTCTTGGTTGGTAAAG 1620
Db 1561 GGTGACTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGTCTTGGTTGGTAAAG 1620
QY 1621 AATGGTTTCACTTACCATGGCAAGCTACTGTGCTTAAATGGGCAAAAGTTCTGATTGG 1680
Db 1621 AATGGTTTCACTTACCATGGCAAGCTACTGTGCTTAAATGGGCAAAAGTTCTGATTGG 1680
QY 1681 TCAGAAAACCTGGGCGGAGCCAAAGGTGCCAACTCTTAAATTTACTAGTTTCGGCACGC 1740
Db 1681 TCAGAAAACCTGGGCGGAGCCAAAGGTGCCAACTCTTAAATTTACTAGTTTCGGCACGC 1740
QY 1741 TCACCAATTCACGACACCCGAAAGTAGCGCTCTCAGCCAGAACTATGCACTTAACCTCACTT 1800
Db 1741 TCACCAATTCACGACACCCGAAAGTAGCGCTCTCAGCCAGAACTATGCACTTAACCTCACTT 1800
QY 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAAAATACTCTGTTGG 1860
Db 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAAAATACTCTGTTGG 1860
QY 1861 GGCATCTGCAGAAACCCAGAACTCTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920
Db 1861 GGCATCTGCAGAAACCCAGAACTCTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920
QY 1921 CTGAGCCCACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGTTGCGGAACCG 1980
Db 1921 CTGAGCCCACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGTTGCGGAACCG 1980
QY 1981 TTGAAGAAAGACTTTCAGCGGAGCGCTTGAACCTTGGACTAA 2019
Db 1981 TTGAAGAAAGACTTTCAGCGGAGCGCTTGAACCTTGGACTAA 2019

RESULT 7

AAT15311

ID AAT15311 standard; DNA; 5049 BP.

XX AAT15311;

XX 14-OCT-1996 (first entry)

XX Non-attenuated canine parvovirus CPV-39 passage 5 DNA.

XX Parvovirus; dog; vaccine; CPV; myocardial disease; enteric disease; ds;
XX SS.

XX Canine parvovirus.

XX WO9614088-A1.

XX 17-MAY-1996.

XX 02-NOV-1995; 95WO-US014207.

XX 08-NOV-1994; 94US-00336345.

XX (CORR) CORNELL RES FOUND INC.

XX Parrish CR, Gruenberg A, Carmichael LE;

XX WPI; 1996-251556/25.

XX

PT	Attenuated CPV strains contg. up to 4 mutation (s) relative to control
PT	virus - useful as a veterinary vaccine against CPV disease in animals,
PT	such as wild or domestic dogs.
XX	
PS	Claim 1; Page 21-24; 42pp; English.
XX	
CC	This viral DNA is isolated from a non- attenuated CPV. The DNA is
CC	preferably derived from vbr440. The DNA is cloned into a vector which is
CC	used to transfect a host cell. The vector used is preferably pGEM3 or
CC	pGEM5Z. The host cells to be transfected ar selected from Norden
CC	Laboratory feline kidney cells, mink lung cells, Madin-Darby canine
CC	kidney cells or canine A72 cells
XX	
SQ	Sequence 5049 BP; 1791 A; 812 C; 1029 G; 1417 T; 0 U; 0 Other;
	Query Match 51.5%; Score 1040.4; DB 2; Length 5049;
	Best Local Similarity 71.4%; Pred. No. 1.8e-269;
	Matches 1432; Conservative 0; Mismatches 536; Indels 18; Gaps 4;
Qy	1 ATGCTCGAAATGCTTACTCTGATGAAGTTTTGGGAGCAACCAACTGGTTAAAGGAAAAA 60
Db	273 ATGCTCGCAACCAAGTACTACTGAGGAAGTTATGAGGAGTAATTGGTTAAAGAAACAT 332
Qy	61 AGTAACCAGGAAGTGTTCTCATTTGTTTTTAAAAATGAAAATGTTCAACTGAATGAAAA 120
Db	333 GCAGAAATGAAGCATTTTCGTTTGGTTTTTAAATGTGACACGTCCTCAACTAAATGGAAG 392
Qy	121 GATATCGATGGATAGTTACAAAAAGAGCGTCGAGGAGGACGAGCTGAAATCTTTTCAAA 180
Db	393 GATGTTTCGTCGGAACAACATAACCAACCAATTCAAAATGAAGAGCTAAACATCTTTAAT 452
Qy	181 CGAGGAGCGGAAACTACTTTGGGACCAAGC-----GAGGACATCGGAATGGGAAACCA 234
Db	453 AGAGGACACAAACAGCAATGGATCAAAACGAGAGAGAAATGGACTGGGAATCGAA 512
Qy	235 GTGGATGAAATGACCAAAAAGCAAGTATTCATTTTGGTCTTTTGGTTAAAAAATGTTTA 294
Db	513 GTTGATAGTCTCGCAAAAAGCAAGTACAAACTTTTGATGCATTAATTAATAAAATGTC 572
Qy	295 TTTTGAAGTCTTAACACAAAGAAATATATTTCTGGTGAATGTTAATTCGTTTGTGCAACAT 354
Db	573 TTTTGAAGTCTTTGTTCTTAAAAATATAGAACAAATGAATGTGTTTGGTTTATTCACAT 632
Qy	355 GAATGGGAAAAAGCAACGAGCTGGCACTGCCATGTACTAATTTGAGGAAAGCACTTAGT 414
Db	633 GAATGGGAAAAAGATCAAGCTGCAATGTTCATGTTTACTTTCATAGTAAGAACTTACAA 692
Qy	415 CAAGCTCAAGGAAATGGTGGAGAAAGGCAACTAAATGTTTACTCGAGCAGATGTTGGTA 474
Db	693 CAAGCAACTGTTAAATGGCTACGCAGACAAATGAATATGTATTGGAGTAGATGTTGGTG 752
Qy	475 ACAGCTGTAATGTGCAACTTAACACAGCTGAAGAAATTAACCTAAGAGAAATAGCAGAA 534
Db	753 ACTCTTTGTCGGTAAACTTAACACCACCTGAAAGATTTAAGCTCAGAGAAATTTGCAGAA 812
Qy	535 GACAAATCAGTGGGTACTCTACTTACTATATAGCATAGCAAAACCAAAAAGACTATACC 594
Db	813 GATAGTGAATGGGTGACTATATTAACATACAGACATAGCDAACAAAAAAGACTATGTT 872
Qy	595 AAGTGTGTTCTTTTGGAAACATGATTGCTTACTATTTTTPAACTAAAAAGAAATTAAGC 654
Db	873 AAAATGGTTCATTTTGGAAATATGATAGCATATTACTTTTTTAAACAAAGAAAAAATGTG 932
Qy	655 ACTAGTCCCAAGAGAGCGGAGCTATTTCTTAGCAGTACTCTGGCTGGAAAACTTAAC 714
Db	933 CACAT---GACAAAAGAAAGTGGCTATTTTTTAAAGTACTGATTCGTTGGAAATTTAAC 989
Qy	715 TTTTAAAAAGAGCGGCGGCATCTAGTGAGCAAACTATACACTGATGACATCGGCA 774
Db	990 TTTATGAAGTATCAAGACAGACAAATTTGTGAGCACATTTTACCTGAAACAAATGAACCA 1049
Qy	775 GAAACGGTTGAAACCAACAGTAACCACTGCGCAGGAAACTAAGCGCGCGAGAAATTCAAACT 834

[illegible]

RESULT 8

RESULT 8	
AAT15312	
ID	AAT15312 standard; DNA; 5049 BP.
XX	
XX	AAT15312;
AC	
XX	
DT	14-OCT-1996 (first entry)
XX	
DE	Attenuated canine parvovirus CPV-39 passage 60 DNA.
XX	
KW	Parvovirus; dog; vaccine; CPV; myocardial disease; enteric disease; ds;
KW	ss.
XX	
OS	Canine parvovirus.

XX	
FN	WO9614088-A1.
XX	
PD	17-MAY-1996.
XX	
PX	02-NOV-1995; 95WO-US014207.
XX	
PX	08-NOV-1994; 94US-00336345.
XX	(CORR) CORNELL RES FOUND INC.
XX	
PA	Parrish CR, Gruenberg A, Carmichael LE;
XX	
PI	WFI; 1996-251556/25.
XX	
DR	Attenuated CPV strains contg. up to 4 mutation (s) relative to control virus - useful as a veterinary vaccine against CPV disease in animals,
XX	such as wild or domestic dogs.
PT	
PT	Claim 2; Page 24-27; 42pp; English.
XX	
CC	This viral DNA is isolated from an attenuated CPV. The DNA is preferably derived from VBI440. The DNA is cloned into a vector which is used to transfect a host cell. The vector used is preferably pGEM3Z or pGEM5Z.
CC	The host cells to be transfected ar selected from Norden Laboratory
CC	feline kidney cells, mink lung cells, Madin-Darby canine kidney cells or canine A72 cells
XX	
SQ	Sequence 5049 BP; 1790 A; 814 C; 1029 G; 1416 T; 0 U; 0 Other;
	Query Match 51.5%; Score 1040.4; DB 2; Length 5049;
	Best Local Similarity 71.4%; Pred. No. 1.8e-289;
	Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;
QY	1 ATGGCTCGGAATGCATTACTCTGTAGTATTTTGGGAGCACACCACTGGTTAAAGGAAAAA 60

Db 1350 GCAATTCCTAGACATGTCAAATTTTGTAGATGTCACGGATGGAATGGATTAAAGTTGT 1409
Qy 1135 CATGCTATTGCTGCTGTTTAAACACAGAGGAGCAAAAGAAATACTGTTTATTTTCAT 1194
Db 1410 CAGCTATTAGCATGTGTTTAAATAGACAGGTTGTTAAAGAAATACAGTTCCTTTTCAT 1469
Qy 1195 GGACACGCGACAGCAGCAAACTATTATTGACAAAGCCATAGCAACAGCAGTTGGCAAT 1254
Db 1470 GGACACGCAAGTACAGGAAATCTATCAITGCTCAAGCCATAGCAACAGCTGTGGTAAT 1529
Qy 1255 GTTGGTTGCTATATGATGAGCCAAATGAACTTTCCATTTTAATGACTGTACCAACAGAAC 1314
Db 1530 GTTGGTTGTTAATGATGAGCAAAATGAAATTTTCCATTTTAATGACTGTACCAATAAAAT 1589
Qy 1315 TTGATTTGGGTAGAGAGCTGTTAACTTTGGACAGCAAGTAAACAGTTTAAAGCAAT 1374
Db 1590 TTAATTTGGATGGAAGCTGTTAACTTTGGTCAACAGATTATCAATTTAAAGCAATC 1649
Qy 1375 TGCTCTGGTCAAACTATTGCTATTGATCAAAAGGAAAGGAGCAGCAAAACAGATTGAACCA 1434
Db 1650 TGTTCTGGACAAACATAGAAATGATCAAAAGGTAAGGAAATGAACCA 1709
Qy 1435 ACACGATCATGATGACCAATGAGAACATTAACAGTGTGATGATGAGTGTGCGAAGAA 1494
Db 1710 ACTCCAGTAATATGACAACTAATGAAATATTAACAAATTTGTGAGAAATGGATGTGAAGAA 1769
Qy 1495 AGACCAAGACACACTCAACCAATCAGACAGAGATGCTTAAACATTCATCTAACACATACC 1554
Db 1770 AGACCTGAACATACAAACCAATTAAGACAGAAATGTTGACATTAAGTTAGTATGTAG 1829
Qy 1555 TTGCCTGGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGTGCTGGTTG 1614
Db 1830 CTTCCAGGAGACTTTGGTTGGTTGATAAAGAAATGGCCCTTAAATATGTGTCATGTTA 1889
Qy 1615 GTAAAGATGTTACCAATCTACATGCGCAAGCTACTGTGCTAAATGGGGCAAGTTCTCT 1674
Db 1890 GTTAAACATGTTTGAATCAACCATGGCTAACTATACACATCAITTTGGGGAAGATACCA 1949
Qy 1675 GATTGTCAGAAACTGGGGGAGCCAAAGTGCCAACTCTATAAATTTACTAGGTTG 1734
Db 1950 GAATGGATGAAACTGGGGGAGCCCTAAATACAGAGATTAATATCACACAGTTGTCAG 2009
Qy 1735 GCACGCTCACCATTCAGCACCCGAAAGTACGCTCTCAGCCAGAACATATGCACTAACT 1794
Db 2010 AAAGACT--TAGAGACACAAGCGGCAAGCAATCTCAGAGTCAAGACCAAGTTCTAACT 2066
Qy 1795 CCACTTGCATCGATCTGAGGACCTGGCTTTAGAGCCTTGGAGCACACCAATATCTCT 1854
Db 2067 CCTCTGACTCCGACGTTAGTGGACCTTGCACCTGGAACCGTGGAGTACTCCAGATACGCT 2126
Qy 1855 GTTGGGGCACTGCAGAAACCCAGAACACTGGTGGGAGCTGGTTCCAAAGCCTGCCAAGAT 1914
Db 2127 ATTGCAGAACTGGAATCAACATCAACCAACTTGGCGTTACTCA-----CAAGAC 2180
Qy 1915 GGTCACTGAGCCCACTCTGGTCAGAGATCGAGAGATTTAGAGCGTCTTCGGTGG 1974
Db 2181 GTGCAAGCGAGTCCGACGTTGTCGGAATAGAGGACAGCCTGAGAGCCATCTTACTTCT 2240
Qy 1975 GAACCGTTGAAGAAAGCTTCAGGA 2000
Db 2241 GAACAAATTGGAAGAAATTTTCGAGA 2266

RESULT 9

AAT88324 standard; DNA; 5049 BP.

AC AAT88324;

XX AC

DT 17-OCT-2003 (revised)

DT 21-MAY-1998 (first entry)

XX DE Attenuated canine parvovirus (vBI440) genomic DNA.

XX Canine parvovirus; CPV; attenuation; vBI440; vaccine; dog;
KW feline panleukopenia virus; mink enteritis virus; infection; ds.
XX
OS Canine parvovirus; vBI440 (ATCC VR 2489).

XX Key Location/Qualifiers

FT mutation 59
FT /tag= C
FT /note= "base 59 is G in CPV-39 (passage 5)"
FT mutation 97
FT /tag= d
FT /note= "base 97 is C or T in CPV-39 (passage 5)"
FT CDS 273..2279
FT /tag= a
FT /note= "NS1/NS2 coding region"
FT CDS 2286..4541
FT /tag= b
FT /note= "VP1/VP2 coding region"
FT mutation 4745
FT /tag= e
FT /note= "base 4745 is T in CPV-39 (passage 5)"
FT mutation 4881
FT /tag= f
FT /note= "base 4881 is C in CPV-39 (passage 5)"

XX WO9742972-A1.

PN 20-NOV-1997.

XX 06-MAY-1997; 97WO-US007584.

XX 15-MAY-1996; 96US-00647655.

PR (CORR) CORNELL RES FOUND INC.

XX Parrish CR, Carmichael LE, Gruenberg A;

XX WPI; 1998-008583/01.

XX Canine parvovirus DNA carrying specific attenuating mutation(s) - used as
vaccines for protection against parvovirus and feline pan-leukopenia
virus infections.
Example 8; Page 34-37; 60pp; English.

XX This DNA sequence comprises an attenuated virus genome derived by serial
passaging (60 times) of virulent canine parvovirus (CPV) type 2b isolate
39 in NUPK feline kidney host cells. The attenuated virus is designated
vBI440 (ATCC VR 2489). It contains 4 mutations relative to the sequence
(see AAT88320) of the control (5th passage) wild-type CPV-2b. 2 Mutations
are within the hairpin formed by the 3' terminal palindromic: the mutation
at nucleotide 59 introduces an A into a G-C rich region within the tip of
the hairpin, disrupting the base pairing in one of the 2 small internal
palindromes within that sequence; the thymine at nucleotide 97 is
adjacent to the mismatched bubble (flip-flop) sequence within the
palindrome. The DNA from attenuated CPV strains (see also AAT88321) is
used for the production of infectious molecular DNA clones, which, in
turn, can be transfected into cells to generate master stocks of the
virus. The attenuated viruses can be used in dogs as a vaccine to protect
against CPV disease, or more generally in cats and minks to protect
against feline panleukopenia virus and mink enteritis virus. The vaccines
protect against the currently prevalent CPV-2b type (and all extant
strains of types 2 and 2a), providing a long term immune response.
(Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 5049 BP; 1791 A; 812 C; 1029 G; 1417 T; 0 U; 0 Other;

XX Query Match 51.5%; Score 1040.4; DB 2; Length 5049;
Best Local Similarity 71.4%; Pred. No. 1.8e-269;
Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;

Qy 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGTTTAAGGAAAA 60

Db 273 ATGCTGGCAACAGTATCTAGGAAGTTATGGAGGAGTAATGGTTAAAGAACAT 332
Qy 61 AGTAACGAGAGGTTCTCATTTGTTTTTAAATAAGAAATGTTCAACTGAATGGAAAA 120
Db 333 GCAGAAAAATGAAGCAATTTGTTGTTTTTAAATGTGACAACTGTCACAACTAAATGGAAAG 392
Qy 121 GATATCGGATGGATAGTTACAAAAGAGCTGCAGGAGACGAGCTGAATCTTTTACAA 180
Db 393 GATGTTGCGTGGAAACAACTATGCCAACCAATTCAAAATGAAGAGCTAACATCTTTAAT 452
Qy 181 CGAGGAGCGGAACTACTCTGGGACCAAGC-----GAGGACATGGAATGGGAAACCA 234
Db 453 AGAGGAGCAAAACAGCAATGGATCAAAACCGAAGAGAAATGAGCTGGAAATCGAA 512
Qy 235 GTGATGAATGATCCAAAAAGCAAGTATTCATTTTGAATCTTTGGTTAAAAAATGTTTA 294
Db 513 GTTGATAGTCTCGCCAAAAAGCAAGTACAAACTTTTGTGATCAATTAATTAATAATGTC 572
Qy 295 TTTGAAGTCTTAACACAAAGATATATTTCTGGTGGTATGTTAATTTGGTTTGTGCAACAT 354
Db 573 TTTGAAGTCTTTGTTCTAAAAATATAGAACCAATGAATGTGTTTGGTTTATTCACAT 632
Qy 355 GAATGGGAAAAGACCAAGCTGCGCTGCTGCTACTAATTCGAGGAAAGGACTTTAGT 414
Db 633 GAATGGGAAAAGATCAAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 692
Qy 415 CAAGCTCAAGGAAAATGGTGGGAAAGCAACTAAATGTTTACTGGAGCAGATGGTTGGTA 474
Db 693 CAAGCAACTGTAATGGCTACGACAGACAAATGATATGTTTACTTCAATAGTAAAGAACTTACAA 692
Qy 475 ACAGCTGTAATGTGCAACTAACACAGCTGGAAGAAATTAACCTAAGAGAAATAGCAGAA 534
Db 753 ACTCTTTGTTGCGTAAACTTAACACCACTGAAAGAAATTAAGCTCAGAGAAATTCAGAA 812
Qy 535 GACAATGAGTGTACTCTACTTACTTATAAGCATAGCAACCAACCAAGAAAGTACTACC 594
Db 813 GATAGTAATGGGTGACTATTAACATACAGACATAAGCAAAACAAAAAGAACTATGTT 872
Qy 595 AAGTGTGTTCTTTTGGAAACATGATTCCTTACTATTTTAACTAAAAAGAAATAAGC 654
Db 873 AATATGTTTCATTTTGGAAATATGATGATATTAATTTTAACTAAAAAGAAATAAGTGC 932
Qy 655 ACTAGTCCACCAAGACGAGGCTATTTTCTTAGAGTGAATCTGCTGGTGGAAAACTAAC 714
Db 933 CACAT---GACAAAAGAAAGTGCGTATTTTAAAGTACTGATTCGTTGAAAAATTAAC 989
Qy 715 TTTTAAAGAGCGGCGCCATCTAGTGCAGAACTATACACTGATGACATGCGGCCA 774
Db 990 TTTATGAAGTATCAAGACAGACAAATTTGTCAGCACACTTTTACACTGAACAAATGAAACCA 1049
Qy 775 GAAACGGTTGAAACCAAGTAAACCTGCGCAGGAACTAAGCGCGCAGAAATTCAAACT 834
Db 1050 GAAACGGTTGAAACCAAGTAAACCTGCGCAGGAACTAAGCGCGCAGAAATTCAAACT 1109
Qy 835 AAAAAAGAAAGTTCTATTAATAAAGTACATTTAAAGAGCTGGTGCATTAAGAGAGTAACCTCA 894
Db 1110 AAAAAAGAAAGTCAATCAATGATGTTCTGCGGACTTTGTTAGTAAAGAGTAAACATCA 1169
Qy 895 CCAGAGACTGGATGATGATGAGCGCAGCAGAGTTCATTTGAAATGATGGCTCAACAGGT 954
Db 1170 CCTGAAGACTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1229
Qy 955 GGAGAAAACTGCTGAAAAATACGCTAGAGATTTGTACACTAATCTAGCCAGAACCAAA 1014
Db 1230 GGTGAAAAATCTTTTAAAAAATACCTTTGAAATTTGTACTTTGACTTTTAGCAAGACAAA 1289
Qy 1015 ACAGCAATTTGACTTAATTTTAAAAAGAGTGAACCCAGCAAACTAACCACTTTTCACTG 1074
Db 1290 ACAGCAATTTGAAATTAATCTTGAAGAGCAGATTAATTAATTAATTAATTAATTAAT 1349
Qy 1075 CTTGACACAAGACCTGCGAATTTTGTCTTTTCTGCTGGAATATGTTAAAGTTTGC 1134

Db 1350 GCAAAATCTAGAACATGTCAAAATTTTGTAGATCCAGGATGAATGGATTTAAAGTTGT 1409
Qy 1135 CATGCTATTTGCTGTGTTTAAACACAGAGGAGCAAAAGAAATACTGTTTATTTTCAT 1194
Db 1410 CAGCTATAGCATGTGTTTAAATAGCAAGGTGGTAAAAAGAAATACAGTTCTTTTTCAT 1469
Qy 1195 GGACAGCAGCAGCAGCAAAATCTATTTTGAACAGCCATAGCAGAGTGGCAAT 1254
Db 1470 GGACAGCAGCAGCAGCAAAATCTATCTGCTCAAGCCATAGCAGAGTGGGTAAT 1529
Qy 1255 GTTGGTGTCTATTAATCAGCCAAATGTAAACTTTTCCATTTTAAATGCTGTACCAACAAGAAC 1314
Db 1530 GTTGGTGTCTATTAATCAGCCAAATGTAAATTTTCCATTTTAAATGCTGTACCAACAAGAAC 1589
Qy 1315 TTGATTTGGGTAGAAAGAGCTGGTAACTTTGGACAGCAAGTAAACAGTTTAAAGCAAT 1374
Db 1590 TTAATTTGGATGCAAGAGCTGGTAACTTTGGTCAACAAAGTTAATCAATTTAAAGCAATC 1649
Qy 1375 TGCTCTGGTCAAACTATTCGCAATTTGATCAAAAGGAAAGGAGCAAGCAAGATTGAACCA 1434
Db 1650 TGTCTGGCAAACTAATTAAGATTTGATCAAAAGGAAAGGAGCAAGCAAGATTGAACCA 1709
Qy 1435 ACACCACTCATCATGACCAAAATGAGAACATTAAGTGTGCTAGAAATAGGCTCGAAGAA 1494
Db 1710 ACTCCAGTAAATTAAGCAAACTAATGAATTAACAAATTTGAGAAATTTGAGAA 1769
Qy 1495 AGACCAAGAACACACTCAACCAATCAGACAGCAAGTCTTAAACATTCATTAACACATACC 1554
Db 1770 AGACCTGAACATACACCAACCAATTAAGACAGCAAGTGTGAACATTTAGTATGTAAG 1829
Qy 1555 TTGCTCTGGTACTTTGGTTTGGTTGACAAAATGAATGGCCCATGATTTGCTGGTTG 1614
Db 1830 CTTCCAGGAGACTTTGGTTTGGTTGATAAGAGAAATGGCCTTTAATATATGTGATGGTTA 1889
Qy 1615 GTAAAGAAATGTTTACCACATCTACATGGCAAGCTACTGTGCTAAATGGGCAAGTTCCCT 1674
Db 1890 GTTAAACATGGTTTGAATCAACCATGGCTAATACATCATCTATGAGGAAAGTACCA 1949
Qy 1675 GATGCTCAGAAAACCTGGCGGAGCCAAAGGTGCCAACTCCTATAAAATTTACTAGTTCG 1734
Db 1950 GAATGGATGAAAACCTGGCGGAGCCCTAAAATACAAAGAGGTATAAATTCACCAGGTTC 2009
Qy 1735 GCAGCTCACCATTACGACACCGAAAAGTACGCTCTCAGCCAGCAACTATGCACTAACT 1794
Db 2010 AAAGACT---TAGAGACACAAGCGGCAAGCAATCTCTCAGAGTCAAGACCAAGTTCTAACT 2066
Qy 1795 CCACCTTGGCATCGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAAAATCTCCT 1854
Db 2067 CCTCTGACTCCGACCGTAGTGGACCTTGCACCTGGAACCGTGGAGTACTCCAGATAGCCT 2126
Qy 1855 GTTCCGGGCACTGCAGAAAACCCAGAACACTGGGGAAAGCTGGTTCCAAAGCCTGCCAAGAT 1914
Db 2127 ATTGCAAGAACTGCAAAATCAACAACTCAACCAACTTTGGCGTTACTCA-----CAAAGAC 2180
Qy 1915 GGTCAAACTGAGCCCAACTTTGGTTCAGAGATCGAGGAGGATTTGAGAGGCTGCTCGGTGCG 1974
Db 2181 GTGCAAGCAGTCCGACGCTGGTCCGAAATAGAGGACAGCTGAGAGCCATCTTTACTTCT 2240
Qy 1975 GAACCGTTGAAGAAAGACTTTCAGCA 2000
Db 2241 GAACAAATGGAAGAAATTTTCGAGA 2266

RESULT 10
AAT88321
ID AAT88321 standard; DNA; 5049 BP.
XX
AC AAT88321;
XX AC
XX 21-MAY-1998 (first entry)
XX
DE Attenuated canine parvovirus genomic DNA.
XX

Db 1230 GGTGAAATCTTTTAAAAAATACACTTGAATTTGTACTTTGACTTTTAGCAAGACAAA 1289
Qy 1015 ACAGCATTTGACTTAATTTTAGAAAAAGCTGAAACACGACAAACTAACCACTTTTCACTG 1074
Db 1290 ACAGCATTTGAAATTAATTAATTTAGAAAAAGCAGATAAATACTAACTAACTAACTTGTATCT 1349
Qy 1075 CTTGACACAGAGCCTGCAGAAATTTTGTCTTTTCATGGCTGGAACATATGTTAAAGTTGC 1134
Db 1350 GCAATTTCTAGAACATGTCAAAATTTTGAATTCAGCGATGGAAATGGATTAAAGTTGT 1409
Qy 1135 CATGCTATTTGCTGTGTTTAAACAGACAGAGGAGGAAAAAGAAATACTGTTTATTTTTCAT 1194
Db 1410 CAGCTATAGCATGTGTTTAAATAGACAAGGTGGTAAAGAAATACAGTCTCTTTTCAT 1469
Qy 1195 GGAACGACGACAGCAAGCAAACTTATTTATGCAAGCCATACACAAAGCAGTTGGCAAT 1254
Db 1470 GGACGACGAGTACAGGAAATCTATCTGCTCAAGCCATAGCACAAGCTGTGGGTAAT 1529
Qy 1255 GTTGGTTGCTATTAATGAGCAATGTAACTTTTCCATTTTAATGACTGTACCAACAAGAAC 1314
Db 1530 GTTGGTTGTTTATTAATGAGCAATGTAAATTTTCCATTTTAATGACTGTACCAATAAAT 1589
Qy 1315 TTGATTTGGGTAGAAAGCTGTAACTTTGGACAGCAAGTAAACAGTTTAAAGCCATT 1374
Db 1590 TTAATTTGGATTCGAGAGCTGTAACTTTGCTCAACAAGTAACTCAATTTAAAGCAATC 1649
Qy 1375 TGCTCTGGTCAAACTATTCGATTTGATCAAAAAGGAAAGGAGGAGCAAGCAAGATTTGAACCA 1434
Db 1650 TGTCTCGCAAAACAATAGAAATTTGATCAAAAAGGTAAGGAAGTAAAGCAATTTGAACCA 1709
Qy 1435 ACACGATCATATGACCAACAATAGAACATTTACAGTGTGAGATAGCTGCGGAAGAA 1494
Db 1710 ACTCCAGTAATTTAGCAACTAATTAAGAAATTTAACAATTTGTGAGAAATTTGGATGGAAGA 1769
Qy 1495 AGACCAAGACACACTCAACCAATCAGACAGACAGAAATGCTTAACATTTCACTTAACACATACC 1554
Db 1770 AGACCTGACATACACACCAATAGACAGAGAAATTTGACAAATTTAGTTAGTTAGTTAG 1829
Qy 1555 TTGCTCTGGTGAATTTGGTTGGTTGACAAAATGAATGGCCCAATGATTTGTGCTGGTTG 1614
Db 1830 CTTCCAGGAGACTTTGGTTGGTTGATTAAGAAAGATGGCTTTAATATGTGATGTTA 1889
Qy 1615 GTAAAGATGTTTACAACTACATCTGACAGTCTGCTAAATGGGGCAAAAGTTCTCT 1674
Db 1890 GTTAAACATGGTTTGAATCAACCATGGCTAACTATACACATCATTTGGGCAAAAGTACCA 1949
Qy 1675 GATTGTCAGAAAACCTGGCGGAGCCAAAGGTGCGCAACTCTATAAATTTACTAGTTCTG 1734
Db 1950 GAATGGATGAATACTGGCGGAGCTTAATAACAAGAGGTATAAATTCACGAGTTGC 2009
Qy 1735 GCAGCTCACCAATTCAGCACACGAAAGTACGCTCTCAGCCAGAACTATGCACTAACT 1794
Db 2010 AAAGACT---TAGAGACACAGCGGCAAGCAATCTCTCAGAGTCAAGCCAAAGTTCTAACT 2066
Qy 1795 CCACTTGCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGAGAGCAGCAACCAATATCTCT 1854
Db 2067 CTTCTGACTCCGAGCAGTGTGAGACCTTGCACTGGAACCGGTGAGTACTCCAGATAGCCT 2126
Qy 1855 GTTGGCGGCACTGCAGAAACCCAGACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGAT 1914
Db 2127 ATTGCAAGAACTGCAATACAAATCAACAACTTTGGGCTTACTCA-----CAAGAC 2180
Qy 1915 GGTCAACTGAGCCCAACTTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTCGTTGG 1974
Db 2181 GTGCAAGCGAGTCCGAGGTGTCGGAATAGAGGAGGAGCAGCTGAGAGCCATCTTACTTCT 2240
Qy 1975 GAACTTTGAGAAAGACTTTTCAGCA 2000
Db 2241 GAACAATTTGGAAGAGATTTTCGAGA 2266

RESULT 11
AAT88320

AA88320 standard; DNA; 5049 BP.
AAT88320;
17-OCT-2003 (revised)
21-MAY-1998 (first entry)
Canine parvovirus 39 passage #5 (wild-type).
Canine parvovirus; CPV; attenuation; vBI440; vaccine; dog;
feline panleukopenia virus; mink enteritis virus; infection; ds.
Canine parvovirus; type 2b isolate 39.
Key Location/Qualifiers
CDS 273..2279
FT /*tag= a
FT /note= "NS1/NS2 coding region"
FT 2286..4541
FT /*tag= b
FT /note= "VP1/VP2 coding region"
XX
XX WO9742972-A1.
XX 20-NOV-1997.
XX 06-MAY-1997; 97WO-US007584.
XX 15-MAY-1996; 96US-00647655.
XX (CORR) CORNELL RES FOUND INC.
XX Parrish CR, Carmichael LE, Gruenberg A;
XX WPI; 1998-008583/01.
XX
XX Canine parvovirus DNA carrying specific attenuating mutation(s) - used as
vaccines for protection against parvovirus and feline pan-leukopenia
virus infections.
XX Example 8; Page 37-40; 60pp; English.
XX This DNA sequence comprises the genome of virulent canine parvovirus type
2b (CPV-2b) isolate 39, obtained after 5 serial passages in NLFK feline
kidney host cells. Further passaging has yielded attenuated virus vBI440
(ATCC VR 2489) (see AAT88324) from the 60th passage, and a claimed virus
(see AAT88321) derived from the 65th passage (ATCC 2528). These
respectively contain 4 and 6 mutations in comparison to the virulent 5th
passage virus. The DNA from attenuated CPV-2b strains is used for the
production of infectious molecular DNA clones, which, in turn, can be
transfected into cells to generate master stocks of the virus. The
attenuated viruses can be used in dogs as a vaccine to protect against
CPV disease, or more generally in cats and minks to protect against
feline panleukopenia virus and mink enteritis virus. The vaccines protect
against the currently prevalent CPV-2b type (and all extant strains of
types 2 and 2a), and provide a long term immune response. (Updated on 17-
OCT-2003 to standardise OS field)
XX
SQ Sequence 5049 BP; 1789 A; 813 C; 1030 G; 1417 T; 0 U; 0 Other;
Query Match 51.5%; Score 1038.8; DB 2; Length 5049;
Best Local Similarity 71.3%; Pred. No. 4.9e-269;
Matches 1431; Conservative 0; Mismatches 557; Indels 18; Gaps 4;
Qy 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 60
Db 273 ATGCTGGCAACCAAGTATCTAGGAGGTATGAGGGAGTAAATTTGGTTAAAGAACAT 332
Qy 61 AGTAACCAAGGAAGTGTCTCATTTGTTTAAAAATGAAAAATGTTCAACTGAATGAAAA 120
Db 333 GCAGAAATGAAGCATTTTCGTTGTTTAAATGTGACACGCTCACTAATATGGAAG 392
Qy 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGAGGAGCTGAAATCTTTTACAA 180


```
FT      /*tag= a  
FT      /note= "see AAP40306"  
FT      2107..3522  
FT      /*tag= b  
FT      /note= "see AAP40675"  
XX      W08402847-A.  
XX      02-AUG-1984.  
XX      19-JAN-1984; 84WO-US000063.  
XX      19-JAN-1983; 83US-00459203.  
PR      06-JAN-1984; 84US-00567968.  
XX      (AMGE-) AMGEN.  
XX      Fox GM;  
XX      WPI; 1984-201354/32.  
DR      P-PSDB; AAP40306, AAP40675.  
XX      Polypeptide obtd. by recombinant DNA methods - for vaccination against  
XX      parvovirus infections in man and animals.  
XX      Claim 10; Table II, Page 33-49; 80pp; English.  
XX      The inventors claim an immunologically active polypeptide for the  
XX      development of vaccinal immunity against parvovirus infection. Also  
XX      claimed are DNA sequences wholly or partly duplicative of defined  
XX      sequences. The polypeptides are used in vaccines for conferring  
XX      protection against parvovirus infections in man and animals. (Updated on  
XX      24-OCT-2003 to standardise OS field)  
SQ      Sequence 3524 BP; 1400 A; 719 C; 648 G; 757 T; 0 U; 0 Other;  
  
Query Match      44.6%; Score 901; DB 1; Length 3524;  
Best Local Similarity      69.7%; Pred. No. 6e-232;  
Matches 1296; Conservative 0; Mismatches 545; Indels 18; Gaps 5;  
  
QY      4 GCTGGAAATCCTTACTCTGATGAAGTTTTCGGAGCAACCACTGGTTTAAAGGAAAAAGT 63  
DB      13 GCGGAAACACTTACTCGGAAGAGGTACTAAAGCTTACCACCTGGCTTCAAGATAATGCT 72  
QY      64 AACCAGGAAGTGTCTCATTTGTTTTTAAATAATGAAATGTTCAACTGTAATGGAAGAAT 123  
DB      73 CAAAAGAAGCAATCTCTTATGATTTTAAACACAAAAGTCAATCTAATGGAAGAA 132  
QY      124 ATCGGATGGAATAGTTACAAAAAGAGCTGCAGAGAGCAGAGCTGAAATCTTTACACGA 183  
DB      133 ATTGCTTGGAAATAACTACAAAGAGATACACAGATGCGGAAATGATAAAACCTTACAAAG 192  
QY      184 GGAGCGAARACTCTTGGGACCAAGCGAGCATGGAATGGGAAACCAAGTGGATGAA 243  
DB      193 GGAGCAGAAACATCATGGGACCAAGCGAACAGACATGGAATGGGAATCGACAGC 252  
QY      244 ATGACCAAAAAGCAAGATTAATTTTGTGTTTAAATAATGTTTATTTGAAGTG 303  
DB      253 CTCACAAAAGCGAAGTACTGATTTTGTACTCTCTTGTAAATAATGTTCTTTGAAGGT 312  
QY      304 CTTAAACAAGAAATATTTCTCGGTGATGTTAATTTGTTTGTGCAACATGAATGGGA 363  
DB      313 ATATTGCAAAAGAACCTTAAGTCCAAGTGAAGTCTGCTTCTTACAGCATGAACATGTT 372  
QY      364 AAGACCAAGCTGGCACTGCCATGTAATTTGGAGGAAGGACTTTAGTCAAGCTCAA 423  
DB      373 CAAGATACCTGGCTATCACTGCCATGTACTAGTGGGAAAGGCTTTACAAAGCAATG 432  
QY      424 GGGAAATGGTGGAGAGGCACTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCTGT 483  
DB      433 GGAATATGGTTACGAAGCAATTAACAAATTTATGGAGTAGATGGTTGAATATCAATGC 492  
QY      484 AATGTGCAACTAACACCAGCTGAAAGAAATTTAAACTAAGAGAAATAGCAGNAGAATGAG 543
```

```
Db      493 AAAGTACCTCTAACACCAGTTGAAAGAAATAAAATTAAGGGAATTTAGCAGAGGATGGTGAG 552  
QY      544 TGGGTTACTCTTACTTATTAAGCATAAAGCAAAACCAAAAAAGACTATATACCAAGTGTGTT 603  
Db      553 TGGGTATCGTCTACTAACCCTACACTCACAAACAACTAAAAAACAATATACAAAAATGACT 612  
QY      604 CTTTTTGGAAAAATGATGCTTACTATTTTAACTAAAAAAGAAATATAGCCTAGTCCCA 663  
Db      613 CATTTTGGAAATATGATGCTTACTACTCTTAAATAAAAAAGAAAGCAACT----- 666  
QY      664 CCAAGAGACGGAGCTTATTTCTTAGCAGTGACTCTGGCTGGGAAAACTACTTTTAAAAA 723  
Db      667 GAAAGAGAGCATGGATATATCTCAGCTCAGATTTCTGGCTTCTATGACAAATTTCTTAAAA 726  
QY      724 GAAGCGAGCGCCATCTAGTGAGCAAAATATACACTGATGACATGCGGCAGAAAAACGGTT 783  
Db      727 GAAGCGAGAGACACTTAGTCAGTCACCTATTTTACTGAGCAATATAACCTGAAACTGTG 786  
QY      784 GAAACACAGTAAACCACTGGCGAGGAAACTAAGCGCGGCGAGAAATTCAAACTAAAAAAGAA 843  
Db      787 GAAACCAACGGTTACTACAGCTCAGGAAGTTTCCCGAGGCGAGAAATACAAAACAAAAAAGAA 846  
QY      844 GTTCTTATTAATACTACACTTAAGAGCTGCTGCATTAAGAGACTAACCTCACCAGAGGAC 903  
Db      847 GTAAGCATAAATGCAATAAGAGACTTGGTTAATAAAGATGTACTAGCATAGAAGGC 906  
QY      904 TGATGATGATGCGAGCGACAGTTACATTTGAAATGATGGCTCAACCCAGCTGGAGAAAAAC 963  
Db      907 TGGAGTATGACAGATCCAGACAGTTATATAGAAATGATGGCTCAACCCGAGGAGAAAT 966  
QY      964 CTGCTGAAAAATAACGCTAGAGATTTGTACACTAACTTAGCCAGAGAACCAAAACAGCATTT 1023  
Db      967 TTAATCAAAAAATACACTAGAAATAACAACCTCTTACTCTAGCAAGAACAAAAACAGCATAT 1026  
QY      1024 GACTTAATTTTAGAAAAAGCTGAAACCAAGCAACTAACCAACTTTTTCAGTGCCTGACACA 1083  
Db      1027 GACTTAATTAATTTGAAAAAGCAAAACCAAGCATGCTACCAACATTTAATATATTAGCAATACA 1086  
QY      1084 AGAGCTGTCAGAAATTTTGTCTTTCATGCTGGAACATATGTTTAAAGTTTGCCTATGCTATT 1143  
Db      1087 AGAACATGTAAATATTTTCAGCATGCAATTTGGAATACATTAAGTCTGCCATGCTATA 1146  
QY      1144 TGCTGTGT-TTTAAACAGACAAGGAGGCAAAAGAAATA-----CTGTTTTATTTCATGGA 1197  
Db      1147 ACTTGTGTACTGAAACAGACAAGGAGGAAAAAGAAATACAAATCTATTTTCTATGTCATGGG 1206  
QY      1198 CAGCCAGCACAGGCAATCTATTATTGACAGCCATAGCAAGCAGTGGCAATGTT 1257  
Db      1207 CCAGCATCAACAGGAAAAAGTAAATTTGCTCAACACATTTGCAAACTTAGTTGTAATGTT 1266  
QY      1258 GGTGCTATAATGACGCAATGTAAACTTTTCCATTTTAACTGTTTACCAAGCAAGAACTTG 1317  
Db      1267 GGTGCTACATGACGCCAATGTGAATTTCCATTTTAACTGTTTACCAAGCAAGAACTTA 1326  
QY      1318 ATTTGGGTAGAAGAAGCTGGTAACTTTTGGACAGCAAGTAAACCAAGTTTAAAGCCATTTGC 1377  
Db      1327 ATATGATTTGAAGAAGCAGGAAACTTCTTAACCAAGTAAACCAATTTCAAAGCCATATGT 1386  
QY      1378 TCTGTTCAAACTATTTCGCAATTGATCAAAAAGGAAAGGCGCAAGCAAGATTTGAACCAACA 1437  
Db      1387 TCAGGTCAAACTATTAGAATTGACCAAAAGGTTAAAGGAGCAAGCAAAATTTGAACCAACT 1446  
QY      1438 CCAGTCATCATGACCAAAATGAGAACATTTACAGTGTGAGAAATAGCTGCGAGAAAGAA 1497  
Db      1447 CTTGTAATTAATGACTAAATGAAGACATACTAAAGTTTGAATAGATGCGAGGAAAGA 1506  
QY      1498 CCAGAACACACTCAACCAATCAGAGACAGAAATGCTTTAACTTCAATCTTAACACATACCTTG 1557  
Db      1507 CCAGAACATACACCAACCAATAGAGACAGAAATGTTAAACATACACCTTAACCAAGAACTG 1566  
QY      1558 CTTGTTGACTTTGTTTGGTTTGAACAAAATGAATGGCCCAATGATTTGTTGCTTGGTTGGTA 1617
```

1567 CCAGGTGATTTGGACCTTTTAGAAGAACTGAATGGCCACTAATATGTCTGGTTGGTA 1626
 1618 AGAATGGTTACCAATCTACCACTGCAAGTCTGCTAAATGGGCAAGTCTCTGAT 1677
 1627 AGAAGAGTTTACCAAGCAATGGCTAGCTATATGTCATCATTTGGGAAATGATCTGAT 1686
 1678 TGCTCAGAAAACCTGGCGGAGCCAAAGGTGCCAATCTCTATATAATTTTACTAGTTTCGGCA 1737
 1687 TGTCTCAGAAAATTTGAGAGGCCAAATATGCTTCCCAATATAATACACCA---ACAGAC 1743
 1738 CGCTCACCATTTCACGACACCGAAAGTACGCTCTTCAGCGACAACTATGCACTACTCCA 1797
 1744 TCTCAGATTTCCACATCAGTGAATACTTGGCCAGCGACATCAATACGACGCAACTCCA 1803
 1798 CTTGCAATCGAATCTCA---GGACCTGGCTTTAGAGCCTTGGAGCACACCAATATCTCC 1853
 1804 ATACAGGAGGACCTGGATTTAGCTTTAGCTTTGGAGCCGTGGAGCGCAACACACC 1862

RESULT 13

ABQ95626
 ID ABQ95626 standard; DNA; 374 BP.

AC ABQ95626;

XX 28-OCT-2002 (first entry)

DT Tumour suppression-related oligonucleotide #1277.

DE Tumour; cytostatic; antiviral; neuroprotective; nontropic; neuroleptic;
 DE tumour suppression; tumour reversion; apoptosis; viral resistance; human;
 DE viral infection; cell degeneration disease; neurodegeneration; ds;
 DE Alzheimer's disease; schizophrenia; immune disease; inflammatory disease.

OS Homo sapiens.

XX FR2819824-A1.

XX 26-JUL-2002.

XX 23-JAN-2001; 2001FR-00000899.

XX 23-JAN-2001; 2001FR-00000899.

XX (MOLE-) MOLECULAR ENGINES LAB SA.

XX Telerman A, Anson R, Tuijnder M, Susini L;

XX WPI; 2002-610803/66.

XX New nucleic acid implicated e.g. in tumor suppression, useful for
 XX diagnosis of tumors, viral infection and cellular degeneration and for
 XX drug screening.

XX Claim 1; Page 360; 623pp; French.

XX The present invention relates to novel human nucleic acid sequences (I).
 XX The present sequence is one such nucleic acid sequence. Expression of (I)
 XX are implicated in tumour suppression or reversion and apoptosis and viral
 XX resistance. (I) are useful as probes or primers for detecting,
 XX identifying, measuring and/or amplifying nucleic acid sequences, as
 XX antisense reagents and for recombinant production of polypeptides. (I),
 XX polypeptides (II) encoded by (I), vector containing (I), cells containing
 XX these vectors and antibodies (Ab) against (II) are all useful for
 XX treatment/prevention of viral, tumour and cell degeneration diseases
 XX (especially neurodegeneration, such as Alzheimer's disease and
 XX schizophrenia). Analysing the expression of (I) is also useful for
 XX diagnosis and/or prognosis of such diseases. Transgenic animals carrying
 XX (I) are used for studying the aetiology of these diseases (also immune
 XX and inflammatory diseases). Note: In the present specification, SEQ ID 1
 XX to 2280 are claimed in Claim 1, however only SEQ ID 1 to 2270 are shown
 XX in the specification

SQ Sequence 374 BP; 116 A; 57 C; 108 G; 84 T; 0 U; 9 Other;
 Query Match 15.2%; Score 307.4; DB 6; Length 374;
 Best Local Similarity 88.2%; Pred. No. 2.5e-72;
 Matches 329; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
 1 TTTTCAACGAGGCGGAGACCACTTGGGACCAAGCGAGGACATGGAATGGGAAACCCAC 233
 174 TTTTCAACGAGGAGCGGAAACTACTTGGGACCAAGCGAGGACATGGAATGGGAAACCCAC 233
 234 AGTGGATGAAATGACCAAAAGCAAGTATTCTTTTGGTCTTTTAAATAATGTTT 293
 61 AGTGGATGCAATGACCAAAAGCAAGTATTCTTTTGGTCTTTTAAATAATGTTT 120
 294 ATTTGAAGTGTCTTAAACAAAGAAATATTTCTGGTGTATGTTTAAATGTTTGTGCAACA 353
 121 GTTTGAGTGTCTAGCACAAGAAATAGCTCTCTAGTAAATGTTTCTTGGTTCGTGCAGCA 180
 354 TGAATGGGAAAGACCAAGGCTGGCACTGCACTTAAATGTTTCTTGGTTCGTGCAGCA 413
 181 TGAATGGGAAAGACCAAGGCTGGCACTGCACTTAAATGTTTCTTGGTTCGTGCAGCA 240
 414 TCAAGCTCAAGGAAATGTTGGAGAGGCAACTAAATGTTTCTTGGTTCGTGCAGCA 473
 241 TCAAGCTCAAGGAAATGTTGGAGAGGCAACTAAATGTTTCTTGGTTCGTGCAGCA 300
 474 AACAGCCTGTATGTGCAACTAAACACAGCTGAAAGAAATTAACCTAGAGAAATAGCAGA 533
 301 GACTGCTGTAATGTGCAACTAAACACAGCTGAAAGAAATTAACCTAGAGAAATAGCAGA 360
 534 AGACATGAGTGG 546
 361 GGACAGTGANNGG 373

RESULT 14

ABQ94779
 ID ABQ94779 standard; DNA; 421 BP.

XX AC ABQ94779;

XX 28-OCT-2002 (first entry)

DE Tumour suppression-related oligonucleotide #430.

XX Tumour; cytostatic; antiviral; neuroprotective; nontropic; neuroleptic;
 XX tumour suppression; tumour reversion; apoptosis; viral resistance; human;
 XX viral infection; cell degeneration disease; neurodegeneration; ds;
 XX Alzheimer's disease; schizophrenia; immune disease; inflammatory disease.

OS Homo sapiens.

XX FR2819824-A1.

XX 26-JUL-2002.

XX 23-JAN-2001; 2001FR-00000899.

XX 23-JAN-2001; 2001FR-00000899.

XX (MOLE-) MOLECULAR ENGINES LAB SA.

XX Telerman A, Anson R, Tuijnder M, Susini L;

XX WPI; 2002-610803/66.

XX New nucleic acid implicated e.g. in tumor suppression, useful for
 XX diagnosis of tumors, viral infection and cellular degeneration and for
 XX drug screening.

XX Claim 1; Page 141; 623pp; French.

XX The present invention relates to novel human nucleic acid sequences (I).

CC The present sequence is one such nucleic acid sequence. Expression of (I)
CC are implicated in tumour suppression or reversion and apoptosis and viral
CC resistance. (I) are useful as probes or primers for detecting,
CC identifying, measuring and/or amplifying nucleic acid sequences, as
CC antisense reagents and for recombinant production of polypeptides. (I),
CC polypeptides (II) encoded by (I), vector containing (I), cells containing
CC these vectors and antibodies (Ab) against (II) are all useful for
CC treatment/prevention of viral, tumour and cell degeneration diseases.
CC (especially neurodegeneration, such as Alzheimer's disease and
CC schizophrenia). Analysing the expression of (I) is also useful for
CC diagnosis and/or prognosis of such diseases. Transgenic animals carrying
CC (I) are used for studying the aetiology of these diseases (also immune
CC and inflammatory diseases). Note: In the present specification, SEQ ID 1
CC to 2280 are claimed in Claim 1, however only SEQ ID 1 to 2270 are shown
CC in the specification
XX
SQ Sequence 421 BP; 125 A; 77 C; 118 G; 96 T; 0 U; 5 Other;

Query Match 14.2%; Score 285.8; DB 6; Length 421;
Best Local Similarity 87.6%; Pred. No. 1.7e-66;
Matches 311; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 173 CTTTACACGAGGCGGAATCTTGGACCAACGAGGACATGGAAATGGAAACCA 232
DB 59 CTTTACACGAGGCGGAGACCACTTGGACCAACGAGGACATGGAAATGGAGGCG 118
QY 233 CAGTGGATGAATGACCAAAAAGCAAGTATTCTTTTGAATCTTTGGTTAAAAAATGTT 292
DB 119 CAGTGGATGACATGACCAAAAAGCAAGTATTCTTTTGAATCTTTGGTTAAAGAGTTT 178
QY 293 TATTGAAAGTCTTAACACAAAGATATTTCTGTGATGATTAATGTTTGTGCAAC 352
DB 179 TGTGTAAGTGCTCAGCACAAAGAACATAGCTCTCTAGTAATGTTACTTGGTTCGTGAGC 238
QY 353 ATGAATGGGAAAGAACCAAGGCTGGCAGTCTGCTACTTAATTTGGAGGAAGGACTTTA 412
DB 239 ATGAATGGGAAAGAACCAAGGCTGGCAGTCTGCTACTTAATTTGGAGGAAGGACTTTA 298
QY 413 GTCAAGCTCAAGGAAATGGTGAGAGGCAACTAAATGTTTACTGGAGCAGATGTTGG 472
DB 299 GTCAAGCTCAAGGAAATGGTGAGAGGCAACTAAATGTTTACTGGAGTATGATGTTGG 358
QY 473 TAACAGCTGTAATGTGCAACTAACACAGCTGAAGAAATTAACCTAAGAGAAAT 527
DB 359 TGACTGCTGTAATGTTCAACTAACACAGCTGAAGAAATTAACCTAAGAGAAAT 413

RESULT 15

AB094724
ID AB094724 standard; DNA; 423 BP.
XX
AC AB094724;
XX
XX
DT 28-OCT-2002 (first entry)
XX
DE Tumour suppression-related oligonucleotide #375.
XX
KW Tumour; cytostatic; antiviral; neuroprotective; nootropic; neuroleptic;
KW tumour suppression; tumour reversion; apoptosis; viral resistance; human;
KW viral infection; cell degeneration disease; neurodegeneration; ds;
KW Alzheimer's disease; schizophrenia; immune disease; inflammatory disease.
XX
OS Homo sapiens.
XX
PN FR2819824-A1.
XX
PD 26-JUL-2002.
XX
XX 23-JAN-2001; 2001FR-00000899.
XX
PR 23-JAN-2001; 2001FR-00000899.
XX
PA (MOLE-) MOLECULAR ENGINES LAB SA.

XX
PI
XX
DR WPI; 2002-610803/66.
XX
PT New nucleic acid implicated e.g. in tumor suppression, useful for
PT diagnosis of tumors, viral infection and cellular degeneration and for
PT drug screening.
XX
PS Claim 1; Page 125-126; 623pp; French.
XX
CC The present invention relates to novel human nucleic acid sequences (I).
CC The present sequence is one such nucleic acid sequence. Expression of (I)
CC are implicated in tumour suppression or reversion and apoptosis and viral
CC resistance. (I) are useful as probes or primers for detecting,
CC identifying, measuring and/or amplifying nucleic acid sequences, as
CC antisense reagents and for recombinant production of polypeptides. (I),
CC polypeptides (II) encoded by (I), vector containing (I), cells containing
CC these vectors and antibodies (Ab) against (II) are all useful for
CC treatment/prevention of viral, tumour and cell degeneration diseases
CC (especially neurodegeneration, such as Alzheimer's disease and
CC schizophrenia). Analysing the expression of (I) is also useful for
CC diagnosis and/or prognosis of such diseases. Transgenic animals carrying
CC (I) are used for studying the aetiology of these diseases (also immune
CC and inflammatory diseases). Note: In the present specification, SEQ ID 1
CC to 2280 are claimed in Claim 1, however only SEQ ID 1 to 2270 are shown
CC in the specification
XX
SQ Sequence 423 BP; 125 A; 74 C; 124 G; 99 T; 0 U; 1 Other;

Query Match 14.1%; Score 284.8; DB 6; Length 423;
Best Local Similarity 89.2%; Pred. No. 3.2e-66;
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
QY 173 CTTTACACGAGGCGGAATCTTGGACCAACGAGGACATGGAAATGGAAACCA 232
DB 71 CTTTACACGAGGCGGAGACCACTTGGACCAACGAGGACATGGAAATGGAGGCG 130
QY 233 CAGTGGATGAATGACCAAAAAGCAAGTATTCTTTTGAATCTTTGGTTAAAAAATGTT 292
DB 131 CAGTGGATGACATGACCAAAAAGCAAGTATTCTTTTGAATCTTTGGTTAAAGAGTTT 190
QY 293 TATTGAAAGTCTTAACACAAAGATATATTTCTGTGATGATTAATGTTTGTGCAAC 352
DB 191 TGTGTAAGTGCTCAGCACAAAGAACATAGCTCTCTAGTAATGTTACTTGGTTCGTGAGC 250
QY 353 ATGAATGGGAAAGAACCAAGGCTGGCAGTCTGCTACTTAATTTGGAGGAAGGACTTTA 412
DB 251 ATGAATGGGAAAGAACCAAGGCTGGCAGTCTGCTACTTAATTTGGAGGAAGGACTTTA 310
QY 413 GTCAAGCTCAAGGAAATGGTGAGAGGCAACTAAATGTTTACTGGAGCAGATGTTGG 472
DB 311 GTCAAGCTCAAGGAAATGGTGAGAGGCAACTAAATGTTTACTGGAGTATGATGTTGG 370
QY 473 TAACAGCTGTAATGTGCAACTAACACAGCTGAAGAAATTAATAA 516
DB 371 TGACTGCTGTAATGTTCAACTAACACAGCTGAAGAAATTAATAA 414

Search completed: January 22, 2005, 11:49:48
Job time : 957.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 19, 2005, 16:54:30 ; Search time 164.5 Seconds
(without alignments)
8723.907 Million cell updates/sec

Title: US-10-069-056-8
Perfect score: 2019
Sequence: 1 agcgtggaatgcttactc.....agccgtgaacttgactaa 2019

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1040.4	51.5	5049	1	US-08-336-345-1
2	1040.4	51.5	5049	1	US-08-336-345-2
3	1040.4	51.5	5049	2	US-08-647-655-1
4	1040.4	51.5	5049	2	US-08-647-655-2
5	109.4	5.4	4680	1	US-08-254-358-1
6	109.4	5.4	4680	1	US-08-475-391-1
7	109.4	5.4	4680	2	US-08-709-609-1
8	109.4	5.4	4680	5	PCT-US95-07178-1
9	109.4	5.4	4681	4	US-09-807-802A-18
10	109.4	5.4	4910	2	US-08-331-384-2
11	109.4	5.4	4910	2	US-08-836-087-2
12	109.4	5.4	4910	3	US-09-246-320-2
13	109.4	5.4	4910	3	US-09-546-738-2
14	109.4	5.4	7214	4	US-09-438-268-1
15	109.4	5.4	7557	4	US-09-770-315-3
16	109.4	5.4	8151	4	US-09-438-268-2
17	109.4	5.4	8179	4	US-09-438-268-5
18	109.4	5.4	8698	4	US-09-770-315-2
19	104.6	5.2	939	4	US-09-532-594B-12
20	104.6	5.2	1197	4	US-09-532-594B-13
21	104.6	5.2	1611	4	US-09-532-594B-14
22	104.6	5.2	1872	4	US-09-532-594B-3
23	104.6	5.2	1872	4	US-09-532-594B-15
24	104.6	5.2	4767	4	US-09-532-594B-1
25	99.8	4.9	969	4	US-09-807-802A-10
26	99.8	4.9	1200	4	US-09-807-802A-8
27	99.8	4.9	1641	4	US-09-807-802A-6

28	99.8	4.9	1872	4	US-09-807-802A-4	Sequence 4, Appli
29	99.8	4.9	4863	4	US-09-807-802A-19	Sequence 19, Appli
30	99.8	4.9	4718	4	US-09-807-802A-1	Sequence 1, Appli
31	64.8	3.2	7218	1	US-08-232-463-14	Sequence 14, Appli
32	61.2	3.0	1141	4	US-09-806-708B-22	Sequence 22, Appli
33	44	2.2	1141	4	US-09-806-708B-22	Sequence 22, Appli
34	40.8	2.0	1664976	4	US-08-916-421B-1	Sequence 1, Appli
35	40.8	2.0	1664976	4	US-09-692-570-1	Sequence 1, Appli
36	40.2	2.0	277	3	US-09-007-005-3	Sequence 3, Appli
37	40.2	2.0	277	3	US-09-244-796-3	Sequence 3, Appli
38	38.6	1.9	116592	4	US-09-818-512-3	Sequence 2813, Ap
39	38.4	1.9	832	4	US-09-621-976-2813	Sequence 14265, A
40	38.4	1.9	1367	4	US-09-270-767-14265	Sequence 1, Appli
41	38.4	1.9	6328	3	US-08-913-832A-1	Sequence 1, Appli
42	38.4	1.9	6328	4	US-09-249-181A-1	Sequence 1, Appli
43	38.4	1.9	6328	4	US-09-158-707-1	Sequence 1, Appli
44	38.4	1.9	6475	4	US-09-620-312D-325	Sequence 325, App
45	37.8	1.9	399	4	US-09-621-976-8976	Sequence 8976, Ap

ALIGNMENTS

RESULT 1
US-08-336-345-1
; Sequence 1, Application US/08336345
; Patent No. 5814510
; GENERAL INFORMATION:
; APPLICANT: Parrish, Colin R.
; APPLICANT: Gruenberg, Allen
; APPLICANT: Carmichael, Leland E.
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,345
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Gordon, Jennifer
; REGISTRATION NUMBER: 30753
; REFERENCE/DOCKET NUMBER: 7937-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Parvovirus
; US-08-336-345-1

Query Match 51.5%; Score 1040.4; DB 1; Length 5049;
Best Local Similarity 71.4%; Pred. No. 7.9e-293;
Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;
QY 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTTAAGGAAAAA 60

Db 273 ATGCTCGCAACCCAGTATATCTGAGGAAGTTATGAGGGAGTAAATTTGGTTAAAGAAACAT 332
 QY 61 AGTAACAGAGAGTGTCTCATTTGTTTAAATGAAATGTTCAACTGGAATGGAAGA 120
 Db 333 GCAGAAATGAAAGCATTTTCGTTGTTTAAATGTGCAAGCTCAACTAAATGGAAG 392
 QY 121 GATATCGGATGGAATAGTTTACAAAAGAGCTGAGGAGGAGCTGAAATCTTTACAA 180
 Db 393 GATGTTGCTGGAAACAACTATACCAACCAATTCAAATGGAAGCTACACTTTAAAT 452
 QY 181 CGAGGAGCGGAAACTACTTTGGGACCAAGC-----GAGGACATGGAATGGGAAACCA 234
 Db 453 AGAGGAGCAAAACAGCAATGGAATCAACCGAAGAGAAAGATGACCTGGGAATCGGAA 512
 QY 235 GTGGATGAATGACCAAAAGCAAGTATTCATTTTGTGTTGTTGTTTAAATAAGTTTAA 294
 Db 513 GTTGATAGTCTCGCCAAAAGCAAGTACAAACTTTTGTGATTAATTAATAAATGCTTT 572
 QY 295 TTTGGAAGTCTTAACACAAAGATATATTTCTGTGTGATGTTAAATTTGTTTGTGCAACAT 354
 Db 573 TTTGAAGTCTTTTCTAAANATATAGAACCAATGAATGTGTTGTTTATTCACAT 632
 QY 355 GAAATGGGAAAGACCAAGCTGGCACTGCACTGTAATTAATTTGGAGGAAGGACTTTAGT 414
 Db 633 GAATGGGAAAGATCAAGGCTGGCATTTGTCTATGTTTACTTATAGTAAGAACTTACAA 692
 QY 415 CAAGCTCAAGGAAATGTTGGAGAGGCACTAAATGTTTACTTGGAGGAGATGTTGGTGA 474
 Db 693 CAAGCAACTGGTAAATGGCTACCGACACAAATGAATGATGTTTGGAGTATGTTGGTG 752
 QY 475 ACAGCTGTAAATGTGCACTAACACACAGCTGAAAGAAATTAACCTAAGAGAAATPAGCAG 534
 Db 753 ACTCTTTGTTGCGTAACTTAAACCAACTGAAAGATTTAAGCTCAGAGAAATTCAGAA 812
 QY 535 GACAAATGAGTGGTACTCTTACTTATTAAGCATAGCAAAACCAAAAGAAAGCTATACC 594
 Db 813 GATAGTGAATGGGTGACTATTAATTAACATACAGACATAAGCAAAACAAAGAAAGCTATG 872
 QY 595 AAGTGTGTTCTTTTGGAAACATGATTGCTTACTATTTTAACTAAAGAAAGAAATTAAGC 654
 Db 873 AAAATGGTTTCAATTTTGGAAATATGATAGCATATTTACTTTTAAAGAAAGAAATTTGTC 932
 QY 655 ACTAGTCCCAAGAGACGGAGGCTATTTTCTTAGCAGTGACTCTGCTGGAAACTAAC 714
 Db 933 CACAT---GACAAAGAAAGTGGCTATTTTAAAGTACTGATTCTGTTGGAATTTAAC 989
 QY 715 TTTTAAAGAGCGGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATCGCGCCA 774
 Db 990 TTTATGAAGTATCAAGACAGACAAATTTGTGAGCACACTTTTACACTGAACAAATGAACCA 1049
 QY 775 GAAACGGTTGAAACCAAGTAAACCACTGCGCAGGAACTAAGCGCGGAGAAATTCAAACT 834
 Db 1050 GAAACCGTTGAAACCAAGTAAACCACTGCGCAGGAACTAAGCGCGGAGAAATTCAAACT 1109
 QY 835 AAAAAAGAGTTTCTATTAAAACTTACACTTTAAAGAGCTGGTGCATAAAAAGAGTAACCTCA 894
 Db 1110 AAAAGGAGTGTCAATCAATGATTTCTTGGGACTTGGTTAGTAAAGAGTAAACATCA 1169
 QY 895 CCAGAGACTGGATGATGATGAGCCAGACAGTTCATTTGAATGATGGCTCAACAGGT 954
 Db 1170 CCTGAAGACTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1229
 QY 955 GGAGAAACCTGCTGAAATAATACCTAGAGATTTGTACACTTAATCTAGCGCAGAACCAAA 1014
 Db 1230 GGTGAAATCTTTTAAATAATACACTTGAATTTGTACTTTGACTTTAGCAAGACAAAA 1289
 QY 1015 ACAGCAATTGACTTTAAATTTTAGAAAAAGCTGAAACCAAGCAAACTAACCACTTTTCACTG 1074
 Db 1290 ACAGCAATTGAAATTAATTTGAAAAAGCAGATAATTAATAACTTAACTTAACTTTGATCTT 1349
 QY 1075 CCTGACACAGAGCCTGCAAGATTTTGTCTTTTCATGCTGGAACTATGTTAAAGTTTGC 1134

Db 1350 GCAAAATCTAGAACCATGTCAAAATTTTAAAGATGACGAGTGAATTTGATTAAGTTTGT 1409
 QY 1135 CATGCTATTTGCTGTGTTTAAACAGACAGAGGAGCAAAAGAAATACGTTTATTTTCAAT 1194
 Db 1410 CACGCTATAGCATGTGTTTAAATAGACAGAGTGTGTAAGAAATACAGATTTCTTTTCAAT 1469
 QY 1195 GGACGAGCAGACAGGCAAAATCTATTTATTTGCAAGCCATAGCAAGAGCTTTGGCAAT 1254
 Db 1470 GGACGAGCAAGTACAGGAAATCTATCTATGCTCAAGCCATAGCAAGAGCTTTGGGTAAAT 1529
 QY 1255 GTTGTGCTGCTAATAATGACGCAATGTAATCTTTCCATTTAATGACTGTACCAACAAAGAAC 1314
 Db 1530 GTTGTGCTGCTAATAATGACGCAATGTAATCTTTCCATTTAATGACTGTACCAATTAATAAT 1589
 QY 1315 TTGATTTGGGTAGAAAGAGCTGGTAACTTTTGACAGCAAGTAAACCAAGCTTTTAAAGCCAT 1374
 Db 1590 TTAATTTGGAATGGAAGAGCTGGTAACTTTGCTCAACAAAGTTAATCAATTTTAAAGCAATC 1649
 QY 1375 TGCTCTGCTCAAACTATTTGCAATTTGATCAAAAGAAAGGAGCAAGCAAGATTTGCAACA 1434
 Db 1650 TGCTCTGCAACAAATTTAGAAATTTGATCAAAAGGTAAGGAAAGTAAAGCAATTTGAACCA 1709
 QY 1435 ACACGAGTCATCATCACCAACCAATGAGAACATTAAGTGTGAGTGTGAGCTTGCAGAA 1494
 Db 1710 ACTCCAGTAATTTAGCAAACTTAATGAAATTAACAAATTTGTGAGAAATTTGATGTGAGAA 1769
 QY 1495 AGACGAGAACACACTCAACCAATCAGACAGAAATGCTTTAAACATTCATCTAACACATACC 1554
 Db 1770 AGACCTGAACATFACACCAACCAATTAAGAGACAGAAATGTTGAACATTTAAGTTAGTATGTAAG 1829
 QY 1555 TTGCTCTGCTGACTTTGTTGTTGTTGCAAAATGATGCCCCATGATTTGCTTTGTTGTTG 1614
 Db 1830 CTTCCAGGAGACTTTGTTGTTGTTGATTAAGAAAGAAATGGCTTTAATATGTCATGGTTA 1889
 QY 1615 GTAAAGATGTTTCAATCTACCTACCTGCAAGCTACTGTGCTAAATGGGCGCAAAAGTTCTCT 1674
 Db 1890 GTTAAACATGTTTGAATCAACCAATGCTAACTATACACATCAATTTGGGAAAGTACCA 1949
 QY 1675 GATGCTCAGAAAACTGGGCGGAGCCAAAGGTGCGCAACTCTTATAAATTTACTAGGTTGCG 1734
 Db 1950 GAATGGGATGAAAACTGGGCGGAGCTTAAATFACAAAGAGTATAAATTTCAACAGTTGC 2009
 QY 1735 GCAGCTCACATTCACGACACGAAAGTACGCTCTCAGCCAGAACTATGCACTAACT 1794
 Db 2010 AAAGACT---TAGAGACACAGCGGCAAGCAATCTCAGATCAGACCAAGCTTTACT 2066
 QY 1795 CCACTTGCAATCGGATCTCGAGGACCTGCTTTAGAGCTTTAGAGCTTTGGAGCACACCAATACTCT 1854
 Db 2067 CTTCTGACTCCGAGCGTAGTGGACCTTGGCACTGGAAACCGTGGAGTACTCCAGATACGCT 2126
 QY 1855 GTTGGGGCACTGCAAGAAACCCAGAACACTGGGGAGCTGGTTCCAAAGCCTGCCAGAT 1914
 Db 2127 ATTGCAGAACTGCAAAATCAACAATCAACCAACTTTGGCGTTACTCTCA-----CAAAGAC 2180
 QY 1915 GGTCACTAGCCCCCACTTTGTTGAGATCGAGGAGGATTTGAGAGCGGTCTTCGGTGGG 1974
 Db 2181 GTGCAAGCAGTCCGAGCTGGTCCGAAATAGAGGAGACCTGAGAGCCATCTTACTTCT 2240
 QY 1975 GAAACCGTTGAAAGAAAGACTTTGAGCGA 2000
 Db 2241 GAACATTTGGAAGAAATTTTCGAGA 2266

RESULT 2
 US-08-336-345-2
 ; Sequence 2, Application US/08336345
 ; Patent No. 5814510
 ; GENERAL INFORMATION:
 ; APPLICANT: Parrish, Colin R.
 ; APPLICANT: Gruenberg, Allen
 ; APPLICANT: Carmichael, Leland E.
 ; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
 ; NUMBER OF SEQUENCES: 2

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Pennie & Edmonds
;; STREET: 1155 Avenue of the Americas
;; CITY: New York
;; STATE: NY
;; ZIP: 10036-2711
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/336,345
;; FILING DATE:
;; CLASSIFICATION: 424
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Gordon, Jennifer
;; REGISTRATION NUMBER: 30753
;; REFERENCE/DOCKET NUMBER: 7937-006
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 790-9090
;; TELEFAX: (212) 869-8864/9741
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 5049 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: both
;; TOPOLOGY: unknown
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; ORIGINAL SOURCE:
;; ORGANISM: Parvovirus
US-08-336-345-2

Query Match 51.5%; Score 1040.4; DB 1; Length 5049;
Best Local Similarity 71.4%; Pred. No. 7.9e-293;
Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;

QY 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTTCGGAGCAACCAACTGGTTTAAAGGAAAA 60
DB 273 ATGCTGGCAACAGTATCTAGGAAAGTTATGGAGGAGTAATTTGGTTTAAAGAAACAT 332
QY 61 AGTAACAGGAAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGAAAA 120
DB 333 GCAGAAAAATCAAGCATTTTCGTTGTTTTTAAATGTGACAACTGCTCAACTTAATGGAAG 392
QY 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGAGGACGAGCTGAAATCTTTACAA 180
DB 393 GATGTTGCTGGGAACTATATACCAACCAATTTCAAAATGAAGAGCTAACATCTTTAAAT 452
QY 181 CGAGGAGCGAAACTACTTCGGACCAAGC-----GAGGACATGGAATGGGAAACCA 234
DB 453 AGAGGAGCACAAACAGCAATGGATCAACCCGAAGAGAAATGAGCTGGGAATCGAA 512
QY 235 GTGGATGAAATGACCAAAAAAGCAAGTATTCATTTTGTATCTTTGTTTAAAAAATGTTTA 294
DB 513 GTTGATAGTCTCGCCAAAAGCAAGTACAACTTTTGATGCATTAATTAATAATGCTTT 572
QY 295 TTTGAAGTGTCTTAACCAAAAGATATATTTCTGGTGAATGTTAAATGGTTTGTGCAACAT 354
DB 573 TTTGAAGTCTTTGTTCTAAAAATATAGAACCAAAATGATGTTTGGTTTATTCACAT 632
QY 355 GAATGGGAAAGACCAAGCTCGCACTGCATGCTACTAATTTGGAGGAAGGACTTTAGT 414
DB 633 GAATGGGAAAGATCAAGCTCGCATTTGCTTTTACTTTCATAGTAAAGAACTACAA 692
QY 415 CAAGCTCAAGGAAATGGTGGAGAGCAACTAAATGTTTACTTGGAGCAGATGTTGGTA 474
DB 693 CAAGCAACTGGTAATGGCTTACGAGACAAATGATATGTTTGGAGTAGATGTTGGTG 752
QY 475 ACAGCCTGTAATGTCAACTTAACACCAAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAA 534
DB 1810 CTTCCAGGAGACTTTGGTTGGTTGATAAAGAGAAATGGCCCTTTAATATGTCATGCTGTTA 1889

753 ACTCTTTGTTGCGTAAACTTAAACCAACTGAAGAAAGATTAAAGCTCAGAGAAATTCGAGAA 812
535 GACAATGAGTGGGTTACTTACTTACTTATTAAGCATAAGCAAAACCAAAAAAGACTATPACC 594
813 GATAGTGAATGGGTGACTATATTAAACATACAGACATAAGCAAAACAAAAAGACTATGTT 872
595 AAGTGTGTTCTTTTTCGAAACATGATTGCTTACTATTATTTTAACTAAAGAAAGAAATTAAGC 654
873 AAAATGGTTTCATTTTGGAAATATGATAGCATATTACTTTTAAACAAAGAAAAAATTTGTC 932
655 ACTAGTCCACCAAGAGACGAGGCTATTCTTCTAGCAGTACTCTGCTGGAAGAACTTAAC 714
933 CACAT---GACAAAGAAAGTGGCTATTTTAAAGTACTGATTCTGTTGGAATTTAAC 989
715 TTTTAAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGATGATGCGGCCA 774
990 TTTATGAAGTATCAAGACAGACAGCAAAATTTGTCAGCACACTTTACACTGAACAAATGAAACCA 1049
775 GAAACGGTTGAAACCAAGTAACTGCGCAGGAAACTTAAGCGCGGAGAAATTTCAACT 834
1050 GAAACCGTTGAAACCAAGTAACTGCGCAGGAAACTTAAGCGCGGAGAAATTTCAACT 1109
835 AAAAAAGAGTCTTCTATTAAAACTACACTTTAAAGAGCTGTCATATAAGAGTAACCTCA 894
1110 AAAAGGAAGTGTCAATCAATGTACTTTGCGGAGCTTGTGTTAGTAAAGAGTAACATCA 1169
895 CCAGAGGACTGGATGATGATGACGACGACAGTTCATTAATGAATGATGGTCAACACAGGT 954
1170 CTTGAAGACTGGATGATGATGACGACGATAGTTATATTGAATGATGGCACAACACGGA 1229
955 GGAGAAAACTGCTGAAAAATAGCTTAGAGATTTGTGACACTAACTCTAGCAGACACAAA 1014
1230 GGTGAAAAATCTTTTAAAAAATAACACTTGAATTTGTACTTTGACTTTAGCAAGAACAAA 1289
1015 ACAGCAATTTGACTTAAATTTTGAAGAAAGCTGAAACGAGCAAACTAACCACTTTTCACTG 1074
1290 ACAGCAATTTGAAATTAATCTTGAAGAAAGCAGATAATACTAACTAACTTTGATCTT 1349
1075 CTTGACACAGAGCTGACAGAAATTTTGTCTTTTTCATGGCTGGAACTATGTTAAAGTTGC 1134
1350 GCAATTTCTAGAACATGTCAAATTTTGAATGACGAGTGAATTTGATTTAAAGTTGT 1409
1135 CATGCTATTGCTGTTTTTAAACAGACAAAGAGGCAAAAGAAATACTGTTTTTATTTTCAT 1194
1410 CAGCTATAGCATGCTTTTAAATAGCAAGGTGGTAAAGAAATACAGTTCTTTTTCAT 1469
1195 GGACCAAGCAGCAGGCAAACTTATTGTCACAGCCATAGCAGACGAGTTGGCAAT 1254
1470 GGACCAAGCAGTACAGGAAATCTATCATTTGCTCAAGCCATAGCAGCAAGCTGTTGGTAA 1529
1255 GTTGGTTGCTATAATGACGCAATGTAAACTTTTCCATTTAATGACTGTACCAACAGAAC 1314
1530 GTTGGTTGTTAATGACGCAATGTAAATTTTCCATTTAATGACTGTACCAATAAAT 1589
1315 TTGAATTTGGGTAGAGAACTGTAATTTTGGACAGCAAGTAAACCAAGTTTAAAGCAAT 1374
1590 TTAATTTGATTGAAGAGCTGGTAATTTTGGTCAACAGTTAATCAATTTAAAGCAATC 1649
1375 TGTCTGTTCAAACTATTCGATTGATCAAAAGAAAGGAGGAGCAAGTAAAGTGAACCA 1434
1650 TGTCTGGAACCAATTAAGATTGATCAAAAGGTTAAAGGAGTAAAGCAATTTGAACCA 1709
1435 ACACCAAGTATCATGACCAACCAATGAGAACATTATACAGTGTGCTAGAAATAGGCTCGGAGAA 1494
1710 ACTCCAGTAATTTATGACAACTAATGAAATATAAATTTGAGAAATTTGAGTGTGAAGAA 1769
1495 AGACCAAGAACACACTCAACCAATCAGAGACAGAAATCTTAAACATTCATCTTAACACATACC 1554
1770 AGACCTGAACATACACCAACCAATAGAGACAGAAATGTTGAACATTAAGTTAGTGTGAAG 1829
1555 TTGCTGTTGACTTTGTTGTTGTTGATAAAGTAAATGGCCCAATGATTTGTTGCTGGTTG 1614
1830 CTTCCAGGAGACTTTGTTGTTGTTGATAAAGAGAAATGGCCCTTTAATATGTCATGCTGTTA 1889


```
QY 1015 ACAGCAATTTGACTTAATTTTAGAAAAAGCTGAAACCAAGCAAACTAACCAACTTTTCACTG 1074
Db 1290 ACAGCAATTTGAAATTAATATCTTGAAGAAAGCAGATAATACTAACTAACTAACTTTGATCTT 1349
QY 1075 CTTGACACAGAGCCTGCGAATTTTCTTTTCATGCTGGAATCTATGTTAAAGTTTGC 1134
Db 1350 GCAAAATTCCTAGAACATGTCAAAATTTTGAATGCGAGGATGAATTTGGAATTAAGTTTGT 1409
QY 1135 CATGCTATTGCTGTTTAAACAGACAGAGGCAAGAAAGAAATCTGTTTAAATTTTCAAT 1194
Db 1410 CAGCTATAGCATGTGTTTAAATAGACAGGTTGTTAAAGAAATACATGTTCTTTTTCAT 1469
QY 1195 GGACCAAGCAGACAGGCAAAATCTATTATGCAAAAGGCAAGCAAGCAAGTGTGCAAT 1254
Db 1470 GGACCAAGCAAGTACAGAAATCTATCAATGCTCAAGCCATAGCAAGCTGTGGTAAAT 1529
QY 1255 GTTGGTGTATATGACGCAATGTAACTTTCCATTTTAATGACTGTACCAACAGAAC 1314
Db 1530 GTTGGTGTATATGACGCAATGTAAATTTTCCATTTTAATGACTGTACCAATAAAT 1589
QY 1315 TTGATTTGGGTAGAAAGCTGTAACTTTTGACAGCAAGTAAACCAAGTTTAAAGCCATT 1374
Db 1590 TTAATTTGATTTGAAGAACTGTAACTTTTGTCAACAAGTTAATCAATTTAAAGCAATC 1649
QY 1375 TGCTCTGGTCAAACTATTGCAATTTGATCAAAAGGAAAGGCAAGCAAAAGCAATTTGAACCA 1434
Db 1650 TGTCTCGGACAACTAATTAAGATTTGATCAAAAGGTAAGGAAGTAAAGCAATTTGAACCA 1709
QY 1435 ACACCAAGTATCATGACCAATGAGACATTTAGACATGCTGAGATAGGCTGCGAAGAA 1494
Db 1710 ACTCCAGTAAATTTATGACAACTAATGAAATATAAACAATTTGTGAGAAATGGATGTGAAGAA 1769
QY 1495 AGACCAAGAACACACTCAACCAATCAGACAGAGAACTTTAACTTCACTAAACATATCC 1554
Db 1770 AGACCTGAACATACACAACTAAGACAGAGAACTTTGACAACTTAAGTTAGTATGTAG 1829
QY 1555 TTGCTCTGGTGAATTTGGTTGTTGACAAAAATGAATGGCCCATGATTTGCTTGGTTG 1614
Db 1830 CTTCCAGGAGACTTTGGTTGGTTGATAAAGAAAGTGGCTTTAATATGTGCTGGTTA 1889
QY 1615 GTAAGAAATGGTTACCAATCTACCATGCGCAAGTACTGTCTAAATGGGGCAAGTTCTCT 1674
Db 1890 GTTAAACATGTTTGAATCAACCATGGCTAACTATACACATCATTTGGGGAAAGTACCA 1949
QY 1675 GATTGCTCAGAAACTGGGCGAGCAAAAGTGGCAACTCTTATAAATTTACTAGTTCTG 1734
Db 1950 GAATGGATGAATACTGGGCGAGCTTAAATACAAAGGTTAATAATTCACAGGTTGC 2009
QY 1735 GCACGCTCACCATTCAACGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTAACT 1794
Db 2010 AAAGACT---TAGAGACACAAGCGGCAAGCAATCTCTCAGAGTCAAGACCAAGTTCTAACT 2066
QY 1795 CCACTTGCATCGGATCTCGAGGACCTGGCTTTAGACCTTTGGAGCAGACCAAAATCTCT 1854
Db 2067 CCTCTGACTCCGAGCAGCTAGTGGACCTTGCATGGAAACCGTGGAGTACTCCAGATACGCT 2126
QY 1855 GTTGGGGCACTGCAGAAACCCAGACACTGGGGAAGCTGTTTCCAAAGCCTGCCAAGAT 1914
Db 2127 ATTGCAGAACTGCAAAATCAACAACTCAACCTTGGGCTTACTCA-----CAAGAC 2180
QY 1915 GGTCAACTGAGCCCAACTTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTGGTGGG 1974
Db 2181 GTGCAAGCAGTCCGAGCTGGTCCGAATAGAGGACAGACTGAGAGCCATCTTTACTTCT 2240
QY 1975 GAACGCTTGAAGAAAGACTTTGACGGA 2000
Db 2241 GAACAAATTTGAAGAAATTTGAGA 2266
```

RESULT 4

US-08-647-655-2

; Sequence 2, Application US/08647655

```
; Patent No. 5885585
; GENERAL INFORMATION:
; APPLICANT: Parrish, Colin R.
; APPLICANT: Gruenberg, Allen
; APPLICANT: Carmichael, Leland E.
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/647,655
; FILING DATE: On Eyen Date Herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gordon, Jennifer
; REGISTRATION NUMBER: 30,753
; REFERENCE/DOCKET NUMBER: 7937-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Parvovirus
; US-08-647-655-2
```

Query Match 51.5%; Score 1040.4; DB 2; Length 5049;
Best Local Similarity 71.4%; Pred. No. 7.9e-293;
Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;

```
QY 1 ATGGCTGGAAGTCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTTAAAGGAAAA 60
Db 273 ATGCTGGCAACCACTATCTAGGAAAGTTATGGAGGAGTAAATTTGGTTAAGAAACAT 332
QY 61 AGTAACCCAGAAAGTGTCTCATTTGTTTAAAAATGAAAAATGTTCAACTGAATGAAAA 120
Db 333 GCAGAAATGAAGCATTTTCGTTTGTAAATGTGACACGTCACCACTAAATGGAAG 392
QY 121 GATATCGGATGAATGATTTACAAAAAGAGCTGCGAGGAGGACGAGCTGAATCTTTACAA 180
Db 393 GATGTTGCTGGAACCACTATACCAACCAATTTCAAAATGAAGAGCTAACATCTTTAAAT 452
QY 181 CGAGGAGCGAACTACTTTGGGACCAAGC-----GAGGACATGGAATGGGAAACACA 234
Db 453 AGAGAGGACCAACACAGCAATGGATCAACCGAAGAAAGAAATGCGCTGGGAATCGAA 512
QY 235 GTGGATGAATGACCAAAAAAGCAAGTATTTCATTTTGTGTTTAAATAATGTTTAA 294
Db 513 GTTGATAGTCTCGCCAAAAAGCAAGTACAACTTTTGATGCAATTAATTAATAATGCTT 572
QY 295 TTTGAAGTCTTAAACACAAAGAAATATTTCTCGTGGTATGTTAAATTTGGTTTGTGCAACAT 354
Db 573 TTTGAAGTCTTTGTTCTTAAAAATATAGAACCAAAATGAATGTGTTTGGTTTATTCAACAT 632
QY 355 GAATGGGAAAGACCAAGCTGGCACTGCCATGTAATAATTCGAGGAAGGACTTTAGT 414
Db 633 GAATGGGAAAGATCAAGGCTGGCACTTGTCTATTTTACTTCAATAGTAAGAACTTACAA 692
```



```

; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31975
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-475-391-1

Query Match          5.4%; Score 109.4; DB 1; Length 4680;
Best Local Similarity 53.7%; Pred: No. 2.8e-21;
Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 1168 GCGAAAGAAYAACTGTTTATTTATTCATGGACACAGCCAGCAGGCAAAATCTATTATTGCA 1227
DB 1293 GCGAAGAGGAACACCATCTGGCTGTTTGGGCCCTGCAACTACCGGAAGACCAACATCGCG 1352
QY 1228 CAAGCCATAGCCACAAGCAGTTGGCAATGTTGGTTGCTATATAATGCAGCCAATGTAAACTTT 1287
DB 1353 GAGGCCATAGCCACACTGTGCCCTTCTACGGGTGCGTAAACTGGACCAATGAGAACTTT 1412
QY 1288 CCAATTTAATGACTGTACCAACAAGAACTTGATTGGTGAAGAAGCTGGTAACTTTTGA 1347
DB 1413 CCCTTCAACGACTGTGTGCACAGATGGTGATCTGTGGGAGGAGGGAGATGACCGCC 1472
QY 1348 CAGCAAGTAAACAGATTTAAAGCCATTTGCTCTGGTCAAACTATTCGACTTGATCAAAA 1407
DB 1473 AAGGTCGTGGAGTCGGCCAAAGCCATTCTCGGAGGAAGCAAGGTGCGTGGACCAAGAA 1532
QY 1408 GGAAGAGGCAGCNAACAGATTGAACCAACACCACTCATCATGACCACAAATGAGAACATT 1467
DB 1533 TGCAGTCTCTCGGCCCCAGATAGACCCGACTCCCGTGATCGTCACTTCAACACCAACATG 1592
QY 1468 ACAGTGGTCAAGATAGGCTGCGAAGAAAGACCAAGAACACACTCAACCAATCAGAGACAGA 1527
DB 1593 TGGCGCGTGATTGACGGGAACCTACAGCACTTCGACACACGAGCGGTTGCAAGACCGG 1652
QY 1528 ATGCTTAAACATTCATTAACACATACCTTGCCCTGGTGACTTTGGTTGGTGAACAAAAT 1587
DB 1653 ATGTTCAAATTTGAATCACTACCCCGCGCTGGATCATGACTTTGGGAAGGTCAACCAAGCAG 1712
QY 1588 GAA 1590
DB 1713 GAA 1715

RESULT 7
US-08-709-609-1
; Sequence 1, Application US/08709609
; Patent No. 5858775
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and
; METHOD OF INVENTION: Methods
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```


CURRENT APPLICATION NUMBER: US/09/807,802A
CURRENT FILING DATE: 2002-02-21
PRIOR APPLICATION NUMBER: US 60/107,114
PRIOR FILING DATE: 1998-11-05
PRIOR APPLICATION NUMBER: PCT/US99/25694
PRIOR FILING DATE: 1999-11-02
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.1
SEQ ID NO 18
LENGTH: 4681
TYPE: DNA
ORGANISM: AAV-2
US-09-807-802A-18

Query Match 5.4%; Score 109.4; DB 4; Length 4681;
Best Local Similarity 53.7%; Pred. No. 2.9e-21;
Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

```
QY 1168 GGCNAAAGAACTACTGTTTATTTTCATGACGAGCCAGCAGCAGGCAAAATCTATTATTGCA 1227
DB 1293 GGCNAAAGAACTACTGTTTATTTTCATGACGAGCCAGCAGCAGGCAAAATCTATTATTGCA 1352
QY 1228 CAAGCCATAGCAACAGCAGTTGGCAATGTTGGTTGCTATATGACGAGCCAGCAGCAGGCAAAATCTATTATTGCA 1287
DB 1353 GAGCCATAGCAACAGCAGTTGGCAATGTTGGTTGCTATATGACGAGCCAGCAGCAGGCAAAATCTATTATTGCA 1412
QY 1288 CCATTTAATGACTGTATACCAAGAACTTGAATTTGGGTAGAAAGCTCGTAACTTTGGA 1347
DB 1413 CCCTTCAACGACTGTGTCACAAGATGTTGCTGTTGGGTAGAAAGCTCGTAACTTTGGA 1472
QY 1348 CAGCAAGTAAACAGTTTAAAGCCATTTGCTCTGGTCAAACTATTGCGCATTTGATCAAAA 1407
DB 1473 AAGGTCGTGGAGTCGCGCAAAAGCCATTTCTGGAGGAAGCAAGTGGCGGTGACCAAGAA 1532
QY 1408 GGAAGAGGAGCAACAGATTTGAACCAACACAGCTCATCATGACCAACAAATGAGAACATT 1467
DB 1533 TCGAAGTCTCGGCCAGATAGACCGACTCCGCTGATGTCACCTCAACACCAATG 1592
QY 1468 ACAGTGGTCAGAAATAGGCTGCGAAGAAAGACAGAAACACACTCAACCAATCAGAGACAGA 1527
DB 1593 TCGCGCGTATTGACGGGAACCTCAACGACTTCGAACACACAGCAGCGTTGCAAGACCGG 1652
QY 1528 ATGCTTAACATTTCATCAACATACCTTGCCTGGTGAATTTGGTTGGTTGATCAAAAAT 1587
DB 1653 ATGTTCAAAATTTGAATCTACCCCGCTCTGGATCATGACTTTGGGAAGGTCCACAGCAG 1712
QY 1588 GAA 1590
DB 1713 GAA 1715
```

RESULT 10

US-08-331-384-2/c
Sequence 2, Application US/08331384
Patent No. 5856152
GENERAL INFORMATION:
APPLICANT: Wilson, James M.
APPLICANT: Kelley, William M.
APPLICANT: Fisher, Krishna J.
TITLE OF INVENTION: Hybrid Adenovirus-AAV Vector and Methods
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr, P.O. Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/331,384
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: UPNG114905A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4910 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
US-08-331-384-2

Query Match 5.4%; Score 109.4; DB 2; Length 4910;
Best Local Similarity 53.7%; Pred. No. 2.9e-21;
Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

```
QY 1168 GGCNAAAGAACTACTGTTTATTTTCATGACGAGCCAGCAGCAGGCAAAATCTATTATTGCA 1227
DB 1544 GGCNAAAGAACTACTGTTTATTTTCATGACGAGCCAGCAGCAGGCAAAATCTATTATTGCA 1485
QY 1228 CAAGCCATAGCAACAGCAGTTGGCAATGTTGGTTGCTATATGACGAGCCAGCAGGCAAAATCTATTATTGCA 1287
DB 1484 GAGGCCATAGCCACACTGTGCGCTTTCTACGGGTGCTAACTGGACCAATGAGAACTTT 1425
QY 1288 CCATTTAATGACTGTATACCAAGAACTTGAATTTGGGTAGAAAGCTCGTAACTTTGGA 1347
DB 1424 CCCTTCAACGACTGTGTCACAAGATGTTGCTGTTGGGTAGAAAGCTCGTAACTTTGGA 1365
QY 1348 CAGCAAGTAAACAGTTTAAAGCCATTTGCTCTGGTCAAACTATTGCGCATTTGATCAAAA 1407
DB 1364 AAGGTCGTGGAGTCGCGCAAAAGCCATTTCTGGAGGAAGCAAGTGGCGGTGACCAAGAA 1305
QY 1408 GGAAGAGGAGCAACAGATTTGAACCAACACAGCTCATCATGACCAACAAATGAGAACATT 1467
DB 1304 TCGAAGTCTCGGCCAGATAGACCGACTCCGCTGATGTCACCTCAACACCAATG 1245
QY 1468 ACAGTGGTCAGAAATAGGCTGCGAAGAAAGACAGAAACACACTCAACCAATCAGAGACAGA 1527
DB 1244 TCGCGCGTATTGACGGGAACCTCAACGACTTCGAACACACAGCAGCGTTGCAAGACCGG 1185
QY 1528 ATGCTTAACATTTCATCAACATACCTTGCCTGGTGAATTTGGTTGGTTGATCAAAAAT 1587
DB 1184 ATGTTCAAAATTTGAATCTACCCCGCTCTGGATCATGACTTTGGGAAGGTCCACAGCAG 1125
QY 1588 GAA 1590
DB 1124 GAA 1122
```

RESULT 11

US-08-836-087-2/c
Sequence 2, Application US/08836087
Patent No. 5871982
GENERAL INFORMATION:
APPLICANT: Trustees of University of Pennsylvania
APPLICANT: Wilson, James M.
APPLICANT: Kelley, William M.
APPLICANT: Fisher, Krishna J.
TITLE OF INVENTION: Hybrid Adenovirus-AAV Vector and
METHODS OF USE THEREOF
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr, PO Box 457

```

; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,087
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/331,384
; FILING DATE: 28-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GNVN.007PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-5818
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4910 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cdna
;
; US-08-836-087-2
;
; Query Match 5.4%; Score 109.4; DB 2; Length 4910;
; Best Local Similarity 53.7%; Pred. No. 2.9e-21;
; Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;
;
; QY 1168 GGCAGAAAGAAATCTGTTTATTTTCATGACAGCCAGACAGGCAAAATCTATTATTGCA 1227
; DB 1544 GGCAGAGGAGAACCACTCTGGCTGTTTGGCCCTGCAACTACCGGGAAGACCAACATCGCG 1485
; QY 1228 CAAGCCATAGCACAAGCAGTTGGCAATGTTGGTTGCTATATATGACGCCAATGTAACTTT 1287
; DB 1484 GAGGCCATAGCCCACTGTGCCCTTCTACGGGTGCGTAAACTGGACCAATGAGAACTTT 1425
; QY 1288 CCATTTAATGACTGACCAAGCAAGAACTTGATTTGGTAGAGAGAGCTGGTAACCTTTGCA 1347
; DB 1424 CCCTTCAACGACTGTGTGCAAGAGATGGTGATCTGGTGGGAGAGGGGAAGATGACCGCC 1365
; QY 1348 CAGCAAGTAAACAGTTTAAAGCCATTTGCTCTGGTCAAACTATTTCGCAATTGATCAAAA 1407
; DB 1364 AAGTCTGGAGTGGCCAAAGCCATTCTCGGAGGAAGCAAGTGGCGGTGGACCAAGAA 1305
; QY 1408 GGAAGAGGAGCAAAACAGATTGAACCAACACAGTCAATGACCAACAATGAGACATT 1467
; DB 1304 TGCAAGTCTCGGCCCCAGATAGACCCGACTCCCGTGATCGTCCCTCCAAACCAACATG 1245
; QY 1468 ACAGTGTGAGATAGCTGCGAGAGAGACCAAGACCACTCAACCAATCAGAGACAGA 1527
; DB 1244 TGGCCGCTGATTGACGGGAACCTCAACGACCTTCGAAACACAGACGCGGTGCAAGACCG 1185
; QY 1528 ATGCTTAAACATTCATCTAAACACATACCTTCCTCGGTGACATTTGGTTTGGTTGACAAAAT 1587
; DB 1184 ATGTTCAATTTGAACTCACCCCGCGTCTGGATCATGACTTTGGGAAGGTTCACCAAGCAG 1125
; QY 1588 GAA 1590
; DB 1124 GAA 1122
;
; RESULT 12
; US-09-246-320-2/c
; Sequence 2, Application US/09246320
; Patent No. 6251677
;
; GENERAL INFORMATION:
; APPLICANT: Trustees of University of Pennsylvania
; APPLICANT: Wilson, James M.
; APPLICANT: Kelley, William M.
; APPLICANT: Fisher, Krishna J.
; TITLE OF INVENTION: Hybrid Adenovirus-AAV Vector and
; TITLE OF INVENTION: Methods of Use Thereof
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, PO Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/246,320
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/836,087
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GNVN.007PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4910 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cdna
;
; US-09-246-320-2
;
; Query Match 5.4%; Score 109.4; DB 3; Length 4910;
; Best Local Similarity 53.7%; Pred. No. 2.9e-21;
; Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;
;
; QY 1168 GGCAGAAAGAAATCTGTTTATTTTCATGACAGCCAGACAGGCAAAATCTATTATTGCA 1227
; DB 1544 GGCAGAGGAGAACCACTCTGGCTGTTTGGCCCTGCAACTACCGGGAAGACCAACATCGCG 1485
; QY 1228 CAAGCCATAGCACAAGCAGTTGGCAATGTTGGTTGCTATATATGACGCCAATGTAACTTT 1287
; DB 1484 GAGGCCATAGCCCACTGTGCCCTTCTACGGGTGCGTAAACTGGACCAATGAGAACTTT 1425
; QY 1288 CCATTTAATGACTGACCAAGCAAGAACTTGATTTGGTAGAGAGAGCTGGTAACCTTTGCA 1347
; DB 1424 CCCTTCAACGACTGTGTGCAAGAGATGGTGATCTGGTGGGAGAGGGGAAGATGACCGCC 1365
; QY 1348 CAGCAAGTAAACAGTTTAAAGCCATTTGCTCTGGTCAAACTATTTCGCAATTGATCAAAA 1407
; DB 1364 AAGTCTGGAGTGGCCAAAGCCATTCTCGGAGGAAGCAAGTGGCGGTGGACCAAGAA 1305
; QY 1408 GGAAGAGGAGCAAAACAGATTGAACCAACACAGTCAATGACCAACAATGAGACATT 1467
; DB 1304 TGCAAGTCTCGGCCCCAGATAGACCCGACTCCCGTGATCGTCCCTCCAAACCAACATG 1245
; QY 1468 ACAGTGTGAGATAGCTGCGAGAGAGACCAAGACCACTCAACCAATCAGAGACAGA 1527
; DB 1244 TGGCCGCTGATTGACGGGAACCTCAACGACCTTCGAAACACAGACGCGGTGCAAGACCG 1185
; QY 1528 ATGCTTAAACATTCATCTAAACACATACCTTCCTCGGTGACATTTGGTTTGGTTGACAAAAT 1587
; DB 1184 ATGTTCAATTTGAACTCACCCCGCGTCTGGATCATGACTTTGGGAAGGTTCACCAAGCAG 1125
; QY 1588 GAA 1590
; DB 1124 GAA 1122
;
; RESULT 12
; US-09-246-320-2/c
; Sequence 2, Application US/09246320
; Patent No. 6251677
;

```

Db 1184 ATGTTCAATTGAACTCACCCCGCTGCTGGATCATGACTTTGGGAAGGTCAACCAAGCAG 1125

QY 1588 GAA 1590
|||

Db 1124 GAA 1122

RESULT 13
US-09-546-738-2/c
; Sequence 2, Application US/09546738
; Patent No. 6387368
; GENERAL INFORMATION:
; APPLICANT: Trustees of University of Pennsylvania
; Wilson, James M.
; Kelley, William M.
; Fisher, Krishna J.
; TITLE OF INVENTION: Hybrid Adenovirus-AAV Vector and
; Methods of Use Thereof
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, PO Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/546, 738
; FILING DATE: 11-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/246,320
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GNVPN.007PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4910 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-546-738-2

Query Match 5.4%; Score 109.4; DB 3; Length 4910;
Best Local Similarity 53.7%; Pred. No. 2.9e-21;
Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 1168 GCGAAGAAATGCTTTTATTTATGACAGCCAGCAGGCAATCTATTATGCA 1227
|||

Db 1544 GCGAAGAGAACACCATCTGGCTGTTTGGGCTGCAACTACCGGGAAGACCAACATCGCG 1485
|||

QY 1228 CAAGCCATAGCAACAGCAGTTGGCAATGTTGGTTGCTATATGACGCCAATGTAAACTTT 1287
|||

Db 1484 GAGGCCATAGCCACACTGTGCCCTCTACGGGTGGTAACTGACCAATGAGAATTT 1425
|||

QY 1288 CCATTTAATGACTGTACCAACAGAACTTGAATTTGGGTAGAAGAGCTGGTAACTTTGA 1347
|||

Db 1424 CCCTTCAACGACTGTGTGCAAGAAGTGTGATCTGTGGGAGGAGGGAAGATGACCGCC 1365
|||

QY 1348 GAGCAAGTAAACAGTTTAAAGCAATTTGCTCTGGTCAAACTATTTCGCAATTTGATCAAAAA 1407
|||

Db 1364 AAGTCGTGGAGTCGGCAAGCCATTCTCGGAGGAAGAGGTGCCGTGGACCAAGAA 1305
|||

QY 1408 GGAAGAGGAGCAAAACAGATTGTAACCAACACCAAGTCATCATGACCAAAATGAGAACATT 1467
|||

Db 1304 TGCAAGTCTCTGGCCCGAGATAGACCCGACTCCCGTGTGTCACCTCCAACACCAATG 1245
|||

QY 1468 ACAGTGGTCAGAAATAGGCTGCGAAGAAAGACAGACACACTCAACCAATCAGAGACAGA 1527
|||

Db 1244 TGGCCCGTGAATGACGGGAACCTCAACGACCTTCGGAACACAGCAGCGTTGCAAGACCG 1185
|||

QY 1528 ATGCTTAACATTCATCTAACACATACCTTGCCTGGTGACTTTTGGTTTGGTTCACAAAAT 1587
|||

Db 1184 ATGTTCAATTTGAATCTACCCCGCTCTGGATCATGACTTTGGGAAGGTCAACCAAGCAG 1125
|||

QY 1588 GAA 1590
|||

Db 1124 GAA 1122

RESULT 14
US-09-438-268-1/c
; Sequence 1, Application US/09438268
; Patent No. 6491907
; GENERAL INFORMATION:
; APPLICANT: Rabinowitz, Joseph E.
; APPLICANT: Samulski, Richard J.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING
; TITLE OF INVENTION: THE SAME
; FILE REFERENCE: 5470-186
; CURRENT APPLICATION NUMBER: US/09/438,268
; CURRENT FILING DATE: 1999-11-10
; EARLIER APPLICATION NUMBER: 60/107,840
; EARLIER FILING DATE: 1998-11-10
; EARLIER APPLICATION NUMBER: 60/123,651
; EARLIER FILING DATE: 1999-03-10
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 7214
; TYPE: DNA
; ORGANISM: Virus
US-09-438-268-1

Query Match 5.4%; Score 109.4; DB 4; Length 7214;
Best Local Similarity 53.7%; Pred. No. 3.6e-21;
Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 1168 GCGAAGAAATGCTTTTATTTATGACAGCCAGCAGGCAATCTATTATGCA 1227
|||

Db 3899 GCGAAGAGAACACCATCTGGCTGTTTGGGCTGCAACTACCGGGAAGACCAACATCGCG 3840
|||

QY 1228 CAAGCCATAGCAACAGCAGTTGGCAATGTTGGTTGCTATATGACGCCAATGTAAACTTT 1287
|||

Db 3839 GAGGCCATAGCCACACTGTGCCCTCTACGGGTGGTAACTGGAACCAATGAGAATTT 3780
|||

QY 1288 CCATTTAATGACTGTACCAACAGAACTTGAATTTGGGTAGAAGAGCTGGTAACTTTGA 1347
|||

Db 3779 CCCTTCAACGACTGTGTGCAAGAAGTGTGATCTGTGGGAGGAGGGAAGATGACCGCC 3720
|||

QY 1348 GAGCAAGTAAACAGTTTAAAGCAATTTGCTCTGGTCAAACTATTTCGCAATTTGATCAAAAA 1407
|||

Db 3719 AAGGTCGTGGAGTCGGCCAAAGCCATTCTCGGAGGAAGCAAGGTGCGCGTGGACCAAGAA 3660
|||

QY 1408 GGAAGAGGCAACAGAGATTGAACCAACCAAGTCATCATGACCAACCAATGAGACATT 1467
|||

Db 3659 TCAAGTCTCTGGCCCGAGATAGACCCGACTCCCGTGTGATCGTCACTCCCAACCAACATG 3600
|||

QY 1468 ACAGTGGTCAGAAATAGGCTGCGAAGAAAGACAGACACACTCAACCAATCAGAGACAGA 1527
|||

Db 3599 TCGCCCGTGAATGACGGGAACCTCAACGACCTTCGAACACACAGCAGCGGTGCAAGACCG 3540
|||

QY 1528 ATGCTTAACATTCATCTAACACATACCTTGCCTGGTGACTTTTGGTTTGGTTCACAAAAT 1587
|||

Db 3539 ATGTTCAATTTGAACCTACCCGCGTCTGGATCATGACTTTGGGAAGGTCAACCAAGCAG 3480
QY 1588 GAA 1590
Db 3479 GAA 3477

RESULT 15
US-09-770-315-3
; Sequence 3, Application US/09770315
; Patent No. 6429001
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; TITLE OF INVENTION: Recombinant AAV Packaging Systems
; FILE REFERENCE: 20263-501
; CURRENT APPLICATION NUMBER: US/09/770,315
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,536
; PRIOR FILING DATE: 2000-01-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 7557
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: recombinant DNA
US-09-770-315-3

Query Match 5.4%; Score 109.4; DB 4; Length 7557;
Best Local Similarity 53.7%; Pred. No. 3.7e-21;
Matches 227; Conservative 0; Mismatches 196; Indels 0; Gaps 0;
QY 1168 GGCAGAGGAAACACCATCTGGCTGTTGGGCTGCAACTACCGGGAAGACCAACATCGCG 1380
Db 1321 GGCAGAGGAAACACCATCTGGCTGTTGGGCTGCAACTACCGGGAAGACCAACATCGCG 1380
QY 1228 CAAGCCATAGCACCAAGCAGTTGGCAATGTTGGTGTATATGACCCCAATGTAACATTT 1287
Db 1381 GAGGCCATAGCCCACTGTGCTGCTTCTAGGGTGGTAACTGGNCCATGAGACATTT 1440
QY 1288 CCATTTAATGACTGTACCAAGAACTTGAATTTGGGTAGAAGAGCTGGTAACATTTGGA 1347
Db 1441 CCCTTCAACGACTGTGTGCAAGATGGTGTCTCTGGTGGGAGGGAAGATGACCGCC 1500
QY 1348 CAGCAAGTAAACAGTTTAAAGCCATTTGCTCTGGTCAAACTATTGCGATTGATCAAAA 1407
Db 1501 AAGTCTGTGGAGTCGGCCAAAGCCATTTCTCGGAGGAAGCAAGGTGCGGTGGACCCAGAAA 1560
QY 1408 GGAAGAGGAGCAAAACAGATTGAACCAACACAGTCTCATGACCAACAATGAGAACATT 1467
Db 1561 TGCAAGTCTCTGGCCAGATAGACCCGACTCCGCTGATGTCACCTCCAACACCAACATG 1620
QY 1468 ACAGTGGTCAGAAATAGGCTCGGAAGAAAGACCAAGAAACACACTCAACCAATCAGAGACAGA 1527
Db 1621 TGGCCGTGATTGACGGGAACCTCAACGACCTTCCGAACACAGCAGCGCTTGCAAGACCGG 1680
QY 1528 ATGCTTAAACATTCATCTAACACATACCTTGCCTGGTGTGCTTGGTTGGTTGCAAAAAT 1587
Db 1681 ATGTTCAAAATTTGAATCTACCCCGCTCTGGATCATGACTTTGGGAAGGTTCACCAAGCAG 1740
QY 1588 GAA 1590
Db 1741 GAA 1743

Search completed: January 19, 2005, 17:05:49
Job time : 169.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2005, 11:50:09 ; Search time 1072 Seconds
(without alignments)
10821.780 Million cell updates/sec

Title: US-10-069-056-8
Perfect score: 2019
Sequence: 1 atggctggaatgcttactc.....agccgctgaacttgactaa 2019

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications_NA.*
- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
 - 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
 - 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
 - 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
 - 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
 - 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
 - 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
 - 8: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
 - 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
 - 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
 - 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
 - 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
 - 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
 - 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
 - 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
 - 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
 - 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
 - 18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
 - 19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
 - 20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
 - 21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1771	87.7	5121	US-10-647-111A-1	Sequence 1, Appli
2	941.4	46.6	5075	US-10-361-002-3	Sequence 3, Appli
3	941.4	46.6	5075	US-10-361-004-3	Sequence 3, Appli
4	307.4	15.2	374	US-10-466-894-1276	Sequence 1276, Ap
5	285.8	14.2	421	US-10-466-894-430	Sequence 430, App
6	284.8	14.1	423	US-10-466-894-375	Sequence 375, App
7	284.8	14.1	464	US-10-466-894-370	Sequence 370, App
8	284.8	14.1	473	US-10-466-894-367	Sequence 367, App
9	284.8	14.1	486	US-10-466-894-358	Sequence 358, App
10	284.8	14.1	491	US-10-466-894-428	Sequence 428, App
11	284.2	14.1	420	US-10-466-894-438	Sequence 438, App
12	283.8	14.1	343	US-10-466-894-388	Sequence 388, App

13	283.8	14.1	343	18	US-10-466-894-390	Sequence 390, App
14	283.8	14.1	403	18	US-10-466-894-433	Sequence 433, App
15	283.2	14.0	497	18	US-10-466-894-369	Sequence 369, App
16	280.4	13.9	340	18	US-10-466-894-383	Sequence 383, App
17	278.2	13.8	342	18	US-10-466-894-416	Sequence 416, App
18	273.8	13.6	420	18	US-10-466-894-432	Sequence 432, App
19	269.6	13.4	324	18	US-10-466-894-391	Sequence 391, App
20	263.4	13.0	339	18	US-10-466-894-1275	Sequence 1275, Ap
21	261.6	13.0	424	18	US-10-466-894-429	Sequence 429, App
22	163.8	8.1	468	18	US-10-466-894-524	Sequence 524, App
23	162.2	8.0	451	18	US-10-466-894-519	Sequence 519, App
24	162.2	8.0	457	18	US-10-466-894-518	Sequence 518, App
25	162.2	8.0	465	18	US-10-466-894-525	Sequence 525, App
26	162.2	8.0	469	18	US-10-466-894-520	Sequence 520, App
27	162.2	8.0	472	18	US-10-466-894-531	Sequence 531, App
28	162.2	8.0	515	18	US-10-466-894-444	Sequence 444, App
29	162.2	8.0	516	18	US-10-466-894-445	Sequence 445, App
30	162.2	8.0	516	18	US-10-466-894-451	Sequence 451, App
31	162.2	8.0	516	18	US-10-466-894-458	Sequence 458, App
32	162.2	8.0	517	18	US-10-466-894-462	Sequence 462, App
33	162.2	8.0	530	18	US-10-466-894-442	Sequence 442, App
34	162.2	8.0	530	18	US-10-466-894-450	Sequence 450, App
35	161	8.0	516	18	US-10-466-894-447	Sequence 447, App
36	160.6	8.0	530	18	US-10-466-894-456	Sequence 456, App
37	159	7.9	468	18	US-10-466-894-523	Sequence 523, App
38	133.4	6.6	247	18	US-10-466-894-471	Sequence 471, App
39	117	5.8	1386	9	US-09-792-630-32	Sequence 32, Appl
40	117	5.8	1386	10	US-09-953-351-32	Sequence 32, Appl
41	117	5.8	1386	13	US-10-080-376-32	Sequence 38, Appl
42	117	5.8	1386	14	US-10-082-671-38	Sequence 38, Appl
43	117	5.8	1386	14	US-10-097-100-32	Sequence 32, Appl
44	117	5.8	1386	15	US-10-023-208-32	Sequence 32, Appl
45	115.4	5.7	1884	9	US-09-792-630-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1

US-10-647-111A-1
; Sequence 1, Application US/10647111A
; Publication No. US20040209240A1
; GENERAL INFORMATION:
; APPLICANT: ICGO, RICHARD
; APPLICANT: MALERBA, MADDALENA
; TITLE OF INVENTION: ANTI-NEOPLASTIC VIRUSES
; FILE REFERENCE: 604-691
; CURRENT APPLICATION NUMBER: US/10/647,111A
; CURRENT FILING DATE: 2003-08-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 1
; LENGTH: 5121
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Parvovirus H1
; OTHER INFORMATION: with promoter P4 and left hairpin from WVM (phH1)
US-10-647-111A-1

Query Match	87.7%	Score	1771;	DB	18;	Length	5121;
Best Local Similarity	92.3%	Pred. No.	0;				
Matches	1864;	Conservative	0;	Mismatches	155;	Indels	0;
Gaps	0;						
QY	1	ATGGCTGGAATGCTTACTCTGATGAACCTTTGGGAGCAACCACTGGTTAAAGGAAAA	60				
DB	282	ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA	341				
QY	61	AGTAACACGAGAGTGTCTTCATTTGTTTTAAAAATGAAATGTTCAACTGAATGAAAA	120				
DB	342	AGTAACACGAGAGTGTCTTCATTTGTTTTAAAAATGAAATGTTCAACTGAATGAAAA	401				
QY	121	GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTTACAA	180				

Db 402 GATATCGGATGGAATAGTTTACAAAAGAGAGCTGCGAGGAGCGAGCTGAATCTTTACAA 461
Qy 181 CGAGGAGCGGAACCTACTTTGGGACCAAGCGAGGACATGGAATGGGAAACACACAGTGGAT 240
Db 462 CGAGGAGCGGAACCTACTTTGGGACCAAGCGAGGACATGGAATGGGAAACACACAGTGGAT 521
Qy 241 GAAATGACCAAAAGCAAGTATTTCTTTGATCTTTGGTTTAAATAATGTTTATTGAA 300
Db 522 GAAATGACCAAAAGCAAGTATTTCTTTGATCTTTGGTTTAAATAATGTTTATTGAA 581
Qy 301 GTGCTTAAACAAAGATATTTCTTGGTGATGTTTAAATGTTTGGTCAACATGAATGG 360
Db 582 GTGCTTAAACAAAGATATTTCTTGGTGATGTTTAAATGTTTGGTCAACATGAATGG 641
Qy 361 GGAAGAGCAAAAGGCTGGCACTGCCATGTAATTTGGAGGAAAGCACTTTAGTCAAGCT 420
Db 642 GGAAGAGCAAAAGGCTGGCACTGCCATGTAATTTGGAGGAAAGCACTTTAGTCAAGCT 701
Qy 421 CAAGGAAATGTTGGAGAAAGGCACTTAATGTTTACTGGAGCAGATGTTGGTGAACAGCC 480
Db 702 CAAGGAAATGTTGGAGAAAGGCACTTAATGTTTACTGGAGCAGATGTTGGTGAACAGCC 761
Qy 481 TGTAAATGTCACCTAACACAGCTGAAGAAATTAATTAAGAGAAATAGCAGAGACAAAT 540
Db 762 TGTAAATGTCACCTAACACAGCTGAAGAAATTAATTAAGAGAAATAGCAGAGACAAAT 821
Qy 541 GAGTGGGTTACTCTACTTACTTATAAGCATAGCAAAACCAAAAGCACTATACCAAGTGT 600
Db 822 GAGTGGGTTACTCTACTTACTTATAAGCATAGCAAAACCAAAAGCACTATACCAAGTGT 881
Qy 601 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAGAGAAATAAGCACTAGT 660
Db 882 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAGAGAAATAAGCACTAGT 941
Qy 661 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTACTCTGGCTGGGAAAACTAACTTTTAA 720
Db 942 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTACTCTGGCTGGGAAAACTAACTTTTAA 1001
Qy 721 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG 780
Db 1002 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATGATGATGATGATGATGATGATGATG 1061
Qy 781 GTTGAAACCAAGTAAACCACTGCGCAGGAAACTAAGCGCGCAGAAATTCAACTAAAAAA 840
Db 1062 GTGAGAACCAAGTAAACCACTGCGCAGGAAACTAAGCGCGCAGAAATTCAACTAGAGAG 1121
Qy 841 GAAGTTTCTATTAAAACTACCTTAAAGAGCTGGTGCATTAAGAGTAACTCACCAGAG 900
Db 1122 GAGTCTCGATTAAAAACCACTCAAGAGTTGGTACATATAAGAGAGTAACTCACCAGAA 1181
Qy 901 GACTGGATGATGCGAGCGCAGAGCTTACATTTGAAATGATGCTCAACAGGTTGGAGAA 960
Db 1182 GACTGGATGATGCGAGCGCAGAGCTTACATTTGAAATGATGCTCAACAGGTTGGAGAA 1241
Qy 961 AACCTGCTGAAAAATACGCTAGAGATTGTACATACTCTAGCCAGAAACAAAGACGA 1020
Db 1242 AACCTGCTTAAAAATACATAGATCTGTACACTGACTCTAGCAGAAACAAAGACGCC 1301
Qy 1021 TTTGACTTAATTTTAAAAAGCTTGAACACAGCAAACTAACCAACTTTTCACTCCCTGAC 1080
Db 1302 TTTGACTTGAATTTGAAAAAGCTTGAACCAAGCAAACTAGCCAACTTTTCACTGGCTAGC 1361
Qy 1081 ACAAGAGCTGAGAAATTTTGTCTTTTTCATGCTGGAATGTTTAAAGTTTGGCATGCT 1140
Db 1362 ACCAGAACTGTAGAAATCTTGTCTGAGCATGCTGGAATGTTTAAAGTTTGGCATGCT 1421
Qy 1141 APTTGTCTGTTTAAAAAGAGAGGCGCAAAAGAAATCTGTTTATTTTATTTGAGGACCA 1200
Db 1422 ATCTGTGTGTCTGAAATAGACAGAGGCGCAAAAGGAACTGTGCTCTTTTACAGGACCA 1481
Qy 1201 GCCAGCAGGCAATCTATTATTTCACAGCCATAGCAGAGTGGCAATGTTGGT 1260

1482 GCCAGCAGGCAATCTATTATTGCAAGCAATAGCAAGAGCTTGGTAAATGTTGGT 1541
Qy 1261 TGTCTAATTCAGCCCAATGTAAACTTTTCCATTTAAATGACTGTACCAACAGAACTTTGAT 1320
Db 1542 TGTCTAATTCAGCCCAATGTAAACTTTTCCATTTAAATGACTGTACCAACAAACTTTGAT 1601
Qy 1321 TGGGTAGAGAAAGCTGGTAACTTTTGGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT 1380
Db 1602 TGGGTAGAGAAAGCTGGTAACTTTTGGCCAGCAAGTAAACCAATTTCAAGCTATTTGTTCT 1661
Qy 1381 GGTCAAACTTTTGGCATTTGATCAAAAAGGAAAGGAGGAGCAAGTAAACCAACACCA 1440
Db 1662 GGGCAAAACCATACGCAATTTGATCAAAAAGGAAAGGAGGAGCAAGTAAACCAACACCA 1721
Qy 1441 GTCAATCATGACCAAAATGAGAACATTTACAGTGGTCAAGATAGGCTGCGAAGAAAGACCA 1500
Db 1722 GTTATTATGACCAACCAACGAGAACATTTACCGTGGTTAGAAATAGGCTGTGAGGAAAGACCA 1781
Qy 1501 GAAACACTCAACCAATCAGAGACAGAAATGCTTTAAACATTCATCTAACACATACCTTGCCT 1560
Db 1782 GAAACACTCAACCAATCAGAGACAGAAATGCTTTAAACATTCATCTAACACATACCTTGCCT 1841
Qy 1561 GGTCACTTTTGGTTTGGTGTGACAAAAATGAATGGCCCATGATTTTGTGCTTGGTAAAG 1620
Db 1842 GGTCACTTTTGGTTTGGTGTGATAGCAAGATGGCTCTGATCTGTGCTTGGTAAAG 1901
Qy 1621 AATGGTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGCAAAAGTTCCCTGATTGG 1680
Db 1902 AATGGTTACCAATCTACCATGGCTTGTACTGTCTAAATGGGCAAAAGTTCCCTGATTGG 1961
Qy 1681 TCAGAAACTGGGCGGAGCCAAAGTGCCAACTCTATATAATTTACTAGGTTCCGGCACGC 1740
Db 1962 TCAGAGGACTGGGCGGAGCCGAAAGCTAGACACTCTATATAATTCGCTAGGTTCAATGCGC 2021
Qy 1741 TCACCATTCAGCAGACACCCGAAAGTACGCTCTCAGCCAGAACTATGCACTAACCTCCACTT 1800
Db 2022 TCACCATCTGACTCCGAGAAAGTACGCTCTCAGCCAAACTACGCTCTTACTCCACTT 2081
Qy 1801 GATCGGATCTCGAGGACTGGCTTTAGAGCTTTAGAGCTTTGGAGCACCAAAATACTCTCTGTTGCG 1860
Db 2082 GCATCGGACTTTGGGACTTAGCTCTAGAGCTTTGGAGCACCAAAATACTCTCTGTTGCG 2141
Qy 1861 GGCATCGCAGAAACCCAGAACACTGGGGAAGCTGTTCCAAAGCTGCAAGATGGTCAA 1920
Db 2142 GGCATCTGAGCAGCAGCAAAACACTGGGGAGCTGGTTCCACAGCTGCGCAAGGTGCTCAA 2201
Qy 1921 CTGAGCCCAACTTGTCTAGAGATCGAGGAGATTTGAGAGCTGCTTCGCTGCGGAAACCG 1980
Db 2202 CGGAGCCCAACCTGGTCCGAGATCGAGGCGGATTTGAGAGCTTGTCTTCAGTCAGAAACAG 2261
Qy 1981 TTGAGAAAGACTTTCAGCGAGCGCTGAACTTTGGACTAA 2019
Db 2262 TTGAGAGCGCACTTCAACGAGGAGCTGACCTTTGGACTAA 2300

RESULT 2
US-10-361-002-3
; Sequence 3, Application US/10361002
; Publication No. US20040170954A1
; GENERAL INFORMATION:
; APPLICANT: Clearant, Inc.
; APPLICANT: McKenney, Keith
; APPLICANT: Gillmeister, Lidja
; APPLICANT: Marlowe, Kristina
; APPLICANT: Amistead, David
; TITLE OF INVENTION: Pathogen Inactivation Assay
; FILE REFERENCE: CI-0043
; CURRENT APPLICATION NUMBER: US/10/361,002
; CURRENT FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 3
; LENGTH: 5075

; TYPE: DNA
; ORGANISM: Porcine parvovirus
US-10-361-002-3

Query Match 46.6%; Score 941.4; DB 17; Length 5075;
Best Local Similarity 70.4%; Pred. No. 2.1e-237;
Matches 1305; Conservative 0; Mismatches 536; Indels 12; Gaps 3;

4	QY	GCTGGAAAATGCTTACTCTGATGAAGTTTTGGGAGCAACCAACTGGTTAAAGGAAAAAAGT	63
298	Db	GCGGGAAAACACTTACTCGGAAGAGGTAATAAAGAGCTACCAACTGGCTTCAAGATAATGCT	357
64	QY	AACACGAGAGCTGTTCTCATTTGTTTTTAAAAATGTTCAACTGAAATGGAAAAAGAT	123
358	Db	CAAAAAGAGCATTTCTTTATGTTTTTAAAAACAATAAAGTCAATCTTAATGGAAAAAGAA	417
124	QY	ATCCGATGGAAATAGTTACAAAAAGAGCTCGAGGAGCAGAGCTGAAATCTTTTACAACGA	183
418	Db	ATTGCTTGGAAATAACTACAAACAAAGATACCAACAGATCGGAAATGATAAACCTTACAAAGA	477
184	QY	GGAGCGGAAACTACTTGGGACCAAGCGAGGACATGGAATGGGAAACCAAGTGGATGAA	243
478	Db	GGAGCAGAAAACATCATGGGACGAGGCAACAGACATGGAATGGGAATCAGAAATCGACAGC	537
244	QY	ATGACCAAAAAGCAAGTATTTCATTTTTGATTTCTTTGTTAAAAAATCTTTATTTGAAGTG	303
538	Db	CTCACAAAACGGCAGTACTGATTTTTTGACTCTCTTTGTTAAAAAATGCTCTTTGAAGGT	597
304	QY	CTTTAACAACAAGAAATATATTTCCTGGTGATGTTTAATTGGTTTGTGCAACATGAATGGGGA	363
598	Db	ATATTGCCAAAAGAACCTTAAGTCCAAGTGACTGCTACTTGGTTTACACGATGAACATGGT	657
364	QY	AAAGACCAAGGCTGGCACTGCCATGTAATAATTGGAGGAAAAGNCTTTAGTCAAGCTCAA	423
658	Db	CAAGATACTGGCTTACATCGCCATGTACTACTAGGTGGAAAAAGGCTTACAAACAGCAATG	717
424	QY	GGCAAAATGGTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGGTTGGTTAAACAGCCTGT	483
718	Db	GGAAAATGGTTCAGAAAACAATTAACAATTTATGGATGATGTTTAATATGCAATGC	777
484	QY	AATGTGCCAATAACACAGCTGAAGAATTTAAACTAAGAGAAAAATAGCAGAAGACAATGAG	543
778	Db	AAAGTACCTTAACACCCAGTTGAAGAAATAAAATTAAGGCAATTAGCAGAGGATGGTGAG	837
544	QY	TGGGTTACTTACTTACTTATAGCATACGCAAAACCAAAAAGACTATACCAAGTGTGTT	603
838	Db	TGGGTATCGCTACTAAACCTTACCTCAACAAACAACTAAAAAACAATATACAAAAATGACT	897
604	QY	CTTTTTGGAAAATGATGTGCTTACTATTTTTTAACTAAAAAGAAAAATAGCACTAGTCCA	663
898	Db	CATTTTTGGAAATATGATTGCTTACTACTTCTTAATAAANAAGNAAGACAACCT-----	951
664	QY	CCAAGAGACGGAGGCTATTTTTCTTAGCAGTGACTCTGGCTGGAAAAACTATCTTTTAAAAA	723
952	Db	GAAAGAGAGCATGGATATTAATCTCAGCTCAGATTTCTGGCTTCATGACAAATTTCTTAAAA	1011
724	QY	GAAGGGCAGCGCCATCTAGTGGACCAACTATACACTGATGACATGCGGCGCAGAAACGGTT	783
1012	Db	GAAGGGCAGAGACACTTAGTTCAGTCACTTATTACTTGAAGCAATAAACCTGAAACCTGTG	1071
784	QY	GAAACCAACAGTAACCACTGCGCAGGAAAATAAGCGCGCAGAAATTCAAACTAAAAAAGAA	843
1072	Db	GAAACACGGTTTACTACAGCTCAGGAAGCCAAAGAGGCGAGATACAAAACAAAAAAGAA	1131
844	QY	GTTTCTTATTAATACTACACTTAAAGAGCTGGTGCATAAAAAGAGTAACTCACAGAGGAC	903
1132	Db	GTAAAGCATAAAAATGCAAAATAAGAGACTTGGTTAAATAAAAAAGATGCTAGCATAGAAGAC	1191
904	QY	TGATGATGATGACGACGACAGTTCATTTGAATGATGCTCAACAGGTGGAGAAAC	963
1192	Db	TGGATGATGACAGATCAGACAGTTATAGAATATGCTCAACCGGAGGAGAAAAAT	1251
964	QY	CTGCTGAAAAATACGGCTAGAGATTTGTACACTAACTCTAGCCAGAAACCAAAACAGCAATTT	1023

RESULT 3

RESULT 3
US-10-361-004-3
; Sequence 3, Application US/10361004
; Publication NO. US20040170981A1
; GENERAL INFORMATION:
; APPLICANT: Clearant, Inc.
; APPLICANT: McKenney, Keith
; APPLICANT: Gillemeister, Lidja
; APPLICANT: Marlowe, Kristina
; APPLICANT: Armstad, David
; TITLE OF INVENTION: Real-Time Polymerase Chain Reaction Using Large T
; FILE REFERENCE: CI-0042
; CURRENT APPLICATION NUMBER: US/10/361.004

```

; CURRENT FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 3
; LENGTH: 5075
; TYPE: DNA
; ORGANISM: Porcine parvovirus
US-10-361-004-3

Query Match      46.6%; Score 941.4; DB 17; Length 5075;
Best Local Similarity 70.4%; Pred. No. 2.1e-237;
Matches 1305; Conservative 0; Mismatches 536; Indels 12; Gaps 3;

QY 4 GCTGGAAATGCTTACTCTGATCAAGTTTGGGAGCAACCAACTGCTTAAAGGAAAAAGT 63
DB 298 GCGGGAACACATTACTTCGGAAGAGTACTAAAGGCTACCACTGCTTCAAGATATGCT 357

QY 64 AACACGGAAGTGTCTCTCAATTTCTTTTAAAAATGAAATGTTTCAACTGAATGGAAGAT 123
DB 358 CAAAAGAGAGCATCTCTTATGATTTTAAAAACAAAAAGTCAATCTAAATGGAAGAA 417

QY 124 ATCGGATGAATAGTTTACAAAAGAGCTGAGGAGGAGCTGAAATCTTTTCAACGA 183
DB 418 ATTGCTTGAATAACTACAAAGATACACAGATGCGGAATGATAAACCTTCAAGA 477

QY 184 GGAGCGGAATCTTTGGGACCAAGCGAGGACATGGAATGGGAAACCAACAGTGGATGA 243
DB 478 GGAGCGGAATCTTTGGGACCAAGCGAGGACATGGAATGGGAAACCAACAGTGGATGA 537

QY 244 ATGACCAAAAGCAAGTATTTCAATTTTGGTGTGTTTAAAAATGTTTATTTGAGTG 303
DB 538 CTCACAAAACGCAAGTACTGATTTTGGTGTGTTTAAAAATGTTTATTTGAGTG 597

QY 304 CTTAACACAAAGATATATTTCTGCTGATGTTTAAATGTTTGTGCAACATGAATGGGA 363
DB 598 ATATTGCAAAAGAACCTTAAGTCAAGTGAATGTTTAAATGTTTGTGCAACATGAATGGGA 657

QY 364 AAAGACCAAGGCTGGCACTGCCATGTAATTTGAGGAAAGGACTTTTAAAGTCAAGTCAA 423
DB 658 CAAGATATGCTGCTATCACTGCCATGTAATTTGAGGAAAGGCTTTTAAAGTCAAGTCAA 717

QY 424 GGAATATGTTGGAAGAGGCAATTAATTTTACTTGAGAGAGATGTTGTTTAAAGCTGT 483
DB 718 GGAATATGTTGGAAGAGGCAATTAATTTTACTTGAGAGAGATGTTGTTTAAAGCTGT 777

QY 484 AATGTCNACTACACAGCTGAAGCAATTAATTTTAAAGCAATTAAGCAATTAAGCAATTAAG 543
DB 778 AAGTACCTCTAACACAGCTGAAGCAATTAATTTTAAAGCAATTAAGCAATTAAGCAATTAAG 837

QY 544 TGGTTTACTCTTACTTATTAAGCATAAGCAACCAACCAACCAACCAACCAACCAACCAAC 603
DB 838 TGGTTTACTCTTACTTATTAAGCATAAGCAACCAACCAACCAACCAACCAACCAACCAAC 897

QY 604 CTTTTTGGAAACATGATGCTTACTATTTTAACTTAAAGCAATTAAGCAATTAAGCAATTAAG 663
DB 898 CAITTTTGGAAATATGATTTGCTTACTTCTTAAATTAAGCAATTAAGCAATTAAGCAATTAAG 951

QY 664 CCAAGAGACGGAGCTATTTCTTACAGTGAATCTGCTGCTGGAATAACTTCTTTTAA 723
DB 952 GAAAGAGAGCATGATATATCTAGCTGAGATCTGCTGCTGGAATAACTTCTTTTAA 1011

QY 724 GAAAGGAGGAGGAGCTATCTAGTGAGCAACTATATACATGATGACATGCGGCGAGAAACGGTT 783
DB 1012 GAAGGCGAGAGACACTTTAGTCAGTCACCTATTTTACTGAAGCAAAATAAACCTGAACTGTG 1071

QY 784 GAAACCAAGTAACTCTGCGAGGAACTAAGCGGCGAGATTTCAACTTAAAGAA 843
DB 1072 GAAACCAAGTAACTCTGCGAGGAACTAAGCGGCGAGATTTCAACTTAAAGAA 1131

QY 844 GTTTCTTATTAACACTACATCTAAAGAGCTGTGATTAAGAGAGTAACTCTCAGCAGAGGAC 903
DB 1132 GTAAGCATTAATTCACATTAAGAGCTGTGATTAAGAGAGTAACTCTCAGCAGAGGAC 1191

```

RESULT 4

US-10-466-894-1276
; Sequence 1276, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius

```

QY 904 TGGATGATGATGCGAGCCAGACAGTTTCAATTTGAATGATGGCTCAACAGCTGGAGAAAC 963
DB 1192 TGGATGATGACAGATCCAGACAGTTTATATAGAAATGATGGCTCAACCGGAGGAAAAAT 1251

QY 964 CTGCTGAAAAATACGCTAGAGATTTGTACACTTAATCTAGCCAGAACCAAAAAACAGATTT 1023
DB 1252 TTAATCAAAAATACACTAGAAATTAACAATCTTCTCTAGCAAGAACAAAAACAGCATAT 1311

QY 1024 GACTTAATTTTAGAAAAAGCTGAAACCCAGCAAACTAAACCAACTTTTCACTGCTGACACA 1083
DB 1312 GACTTAATTTTAGAAAAAGCTGAAACCCAGCAAACTAAACCAACTTTTCACTGCTGACACA 1371

QY 1084 AGAGCTGCGAATTTTGTCTTTTCACTGCTGCGAATCTATGTTAAAGTTTGCCTGCTATT 1143
DB 1372 AGAATCTGTTAAATTTTCACTGCTGCGAATCTATGTTAAAGTTTGCCTGCTATT 1431

QY 1144 TGCTGTGTTTAAACACAGAGGAGGCAAAAGAAATACTGTTTATTTTATTTTATGAGCCAGCC 1203
DB 1432 ACTTGTGTTTAAACACAGAGGAGGCAAAAGAAATACTGTTTATTTTATTTTATGAGCCAGCC 1491

QY 1204 AGCAGGCAAAATCTATTTTTCACAAGCCATAGACACAAGCAGTTGGCAATGTTGTTGC 1263
DB 1492 TCAACAGGAAAAAGCTATTAATTTCTCAACACATTTGCTCAACCACTTAGTTGGTAATTTGTTGC 1551

QY 1264 TATTAATGAGCAATGTTAACTTTTCCATTTTAACTGCTGACCAACCAAGAACTTGAATTTG 1323
DB 1552 TACAATGAGCAATGTTAACTTTTCCATTTTAACTGCTGACCAACCAAGAACTTGAATTTG 1611

QY 1324 GTAGAAGAGCTGTTAACTTTTGGACAGCAAGTAAACCAAGTTTAAAGCCATTTGCTGCT 1383
DB 1612 ATTGAGAGAGGAGAACTTTCTCTTAAACCACTTAAACCACTTAAACCACTTAAACCACT 1671

QY 1384 CAACTATTTCGATTTGATCAAAAAAGGAGGAGCAAAAGAACTGAACTGAACTGAACTGAACT 1443
DB 1672 CAACTATTTCGATTTGATCAAAAAAGGAGGAGCAAAAGAACTGAACTGAACTGAACTGAACT 1731

QY 1444 ATCATGACCAAAATGAGCAATTTACAGTGTGAGATGAGTGTGAGGAGGAGGAGGAGGAG 1503
DB 1732 ATAATGATCTCAAAATGAGCAATTTACAGTGTGAGATGAGTGTGAGGAGGAGGAGGAGGAG 1791

QY 1504 CACACTCAACCAATCAGAGACAGAAATGCTTAACTTCACTTAAACATATACCTGCTGCT 1563
DB 1792 CATACACAACTCAATTAAGAGACAGAAATGTTAAACATTAACATTAACATTAACATTAAC 1851

QY 1564 GACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1623
DB 1852 GACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1911

QY 1624 GGTACCCTTACCATGCAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1683
DB 1912 GGTACCCTTACCATGCAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1971

QY 1684 GAAAACTGGGCGAGCCAAAGGTGCAACTCTCTTAAATTTTACTAGTTTGGCAGCTCA 1743
DB 1972 GAAAACTGGGCGAGCCAAAGGTGCAACTCTCTTAAATTTTACTAGTTTGGCAGCTCA 2028

QY 1744 CCAATTCACGACCAAGGAGTACGCTCTCAGCCAGAACTATGCTACCTACCTACCTACCT 1803
DB 2029 ATTTCACATCTAGTGAAACTTTGCGCAGCGGCAACACTACGCGAGCACTTCAATACAG 2088

QY 1804 TCGGATCTCGA---GGACCTGCTTTTAGAGCTTTGGAGCCGTTGGAGCCGAGCCAAACACC 1853
DB 2089 GAGGACCTGCTGATTTAGCTTTTAGCTTTGGAGCCGTTGGAGCCGAGCCAAACACC 2141

```

```
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; TITLE OF INVENTION: AND THEIR USE AS MEDICINES
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1276
; LENGTH: 374
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 291, 300, 310, 316, 347, 348, 370, 371, 374
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-1276

Query Match      15.2%; Score 307.4; DB 18; Length 374;
Best Local Similarity 88.2%; Pred. No. 1.6e-70;
Matches 329; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 174 TTTTACACGAGGAGCGGAACTACTTTGGGACCAAGCGAGGACATGGAATGGGAAACCCAC 233
Db 1 TTTTACACGAGGCGGAGACCCACTTTGGGACCAAGCGAGGACATGGAATGGGAGAGCGC 60

QY 234 AGTGGATGAATGACCAAAAGCAAGTATTCATTTTGGTAAATGTTT 293
Db 61 AGTGGATGACATGACCAAAAGCAAGTATTTATTTTGGTAAATGTTT 120

QY 294 ATTGGAAGTCTTAACACAAAGATATTTCTCGTGTGATGTTAAATGTTGTCACAA 353
Db 121 GTTTGAAGTGTCTAGCACAAAGACATAGCTCCTAGTAATGTTTACTTGTTCGTGAGCA 180

QY 354 TGAATGGGAAAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGACTTTAG 413
Db 181 TGNATGGGAAAGGACCAAGGCTGGCACTGTCTATGTTGATTTGGAGGCAAGACTTTAG 240

QY 414 TCAAGCTCAAGGAAATGTTGGAGAAAGCAACTAAATGTTTACTGGAGAGATGTTGTT 473
Db 241 TCAACCTCAAGAAATGTTGGAGAAAGGAGCTAAATGTTGAGTANATGTTGTTGN 300

QY 474 ACAGCCTGTATGTCRACTAACACCACTGAAAGTAACTTAAGTAAAGAAATAGCAGA 533
Db 301 GACTGCTGNAATGTTNCAACTAACACCACTGAAAGAAATTAACCTGNNAGAAATAGCAGA 360

QY 534 AGACAAATGAGTGG 546
Db 361 GGACAGTGANNGG 373

RESULT 5
US-10-466-894-430
; Sequence 430, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Teلمان, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1276
; LENGTH: 421
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 422
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-375
```

```
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 430
; LENGTH: 421
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 276, 405, 417, 419, 421
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-430

Query Match      14.2%; Score 285.8; DB 18; Length 421;
Best Local Similarity 87.6%; Pred. No. 8.9e-65;
Matches 311; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGAGCGGAACTACTTTGGGACCAAGCGAGGACATGGAATGGGAAACCA 232
Db 59 CTTTACACGAGGCGGAGACCCACTTTGGGACCAAGCGAGGACATGGAATGGGAGAGCG 118

QY 233 CAGTGGATGAAATGACCAAAAGCAAGTATTCATTTTGGTAAATGTTT 292
Db 119 CAGTGGATGACATGACCAAAAGCAAGTATTTATTTTGGTAAATGTTT 178

QY 293 TATTTGAAGTCTTAAACACAAAGATATTTCTCGTGTGATGTTAAATGTTTGTCAAC 352
Db 179 TGTTTGAAGTCTCAGCACAAAGACATAGCTCCTAGTAATGTTTACTTGTTCGTGAGC 238

QY 353 ATGAATGGGAAAGACCAAGGCTGGCACTGCCATGTACTTAATTTGAGGAAAGGACTTTA 412
Db 239 ATGAATGGGAAAGACCAAGGCTGGCACTGTCTATGTTGATTTGAGGCAAGGACTTTA 298

QY 413 GTCAGCTCAAGGAAATGTTGGAGAAAGCAACTAAATGTTTACTGGAGCAGATGTTGG 472
Db 299 GTCACCTCAAGAAATGTTGGAGAAAGCAAGTAAATGTTGATGATGATGATGTTGG 358

QY 473 TAACAGCCTGTATGTCMACTAACACCAAGTAAAGAAATTAACCTAAGAGAAAT 527
Db 359 TGACTGCTGTATGTTCACTTAACACCAAGTAAAGAAATTAACCTAAGAGAAAT 413

RESULT 6
US-10-466-894-375
; Sequence 375, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Teلمان, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 375
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 422
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-375
```

Query Match 14.1%; Score 284.8; DB 18; Length 423;
Best Local Similarity 89.2%; Pred. No. 1.6e-64;
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGAGCGGAACTACTTGGACCAAGCGAGGACATGGAATGGGAAACCA 232
DB 71 CTTTACACGAGGAGCGGAGACCACTTGGACCAAGCGAGGACATGGAATGGGAGAGCG 130

QY 233 CAGTGGATGAATGACCAAAAGCAAGTATTCAATTTTGGATCTTTTGGTTAAAAATGTT 292
DB 131 CAGTGGATGACATGACCAAAAGCAAGTATTCAATTTTGGATCTTTTGGTTAAAAATGTT 190

QY 293 TATTTTGAAGTCTTAACACAAAGATATATTTCCCTGGTGATGTTAAATGTTTGGTCAAC 352
DB 191 TGTTCGCTGTAATGTTCAACTAACACCAAGCTGAAAGAAATTTAA 516

QY 353 ATGAATGGGAAAAGCAAGCTGGCACTGGCATGTACTAAATGTTTACTGGAGCAGATGTTGG 472
DB 251 ATGAATGGGAAAAGCAAGCTGGCACTGTCTGTCTGATGTTGGAGGCAAGGACTTTA 310

QY 413 GTCAGCTCAAGGAAATGGTGAGAGCAAGCTGGCACTGGCATGTACTAAATGTTTACTGGAGCAGATGTTGG 472
DB 311 GTCAGCTCAAGGAAATGGTGAGAGCAAGCTGGCACTGGCATGTACTAAATGTTTACTGGAGCAGATGTTGG 370

QY 473 TAACAGCTGTAATGTGCAACTAACACCAAGCTGAAAGAAATTTAA 516
DB 371 TGACTGCCCTGTAATGTTCAACTAACACCAAGCTGAAAGAAATTTAA 516

RESULT 7
US-10-466-894-370
; Sequence 370, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telesman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; TITLE OF INVENTION: AND THEIR USE AS MEDICINES
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 370
; LENGTH: 464
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 414, 435, 459, 462, 464
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-370

Query Match 14.1%; Score 284.8; DB 18; Length 464;
Best Local Similarity 89.2%; Pred. No. 1.7e-64;
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGAGCGGAACTACTTGGACCAAGCGAGGACATGGAATGGGAAACCA 232
DB 55 CTTTACACGAGGAGCGGAGACCACTTGGACCAAGCGAGGACATGGAATGGGAGAGCG 114

QY 233 CAGTGGATGAATGACCAAAAGCAAGTATTCAATTTTGGATCTTTTGGTTAAAAATGTT 292
DB 115 CAGTGGATGACATGACCAAAAGCAAGTATTCAATTTTGGATCTTTTGGTTAAAAATGTT 174

QY 293 TATTTGAAGTCTTAAACACAAAGATATATTTCCCTGGTGATGTTAAATGTTTGGTTCGCAAC 352
DB 175 TGTTCGCTGTAATGTTCAACTAACACCAAGCTGAAAGAAATTTAA 516

QY 353 ATGAATGGGAAAAGCAAGCTGGCACTGGCATGTACTAAATGTTTACTGGAGCAGATGTTGG 472
DB 235 ATGAATGGGAAAAGCAAGCTGGCACTGTCTGTCTGATGTTGGAGGCAAGGACTTTA 294

QY 413 GTCAGCTCAAGGAAATGGTGAGAGCAAGCTGGCACTGGCATGTACTAAATGTTTACTGGAGCAGATGTTGG 472
DB 295 GTCAGCTCAAGGAAATGGTGAGAGCAAGCTGGCACTGGCATGTACTAAATGTTTACTGGAGCAGATGTTGG 354

QY 473 TAACAGCTGTAATGTGCAACTAACACCAAGCTGAAAGAAATTTAA 516
DB 355 TGACTGCCCTGTAATGTTCAACTAACACCAAGCTGAAAGAAATTTAA 516

RESULT 8
US-10-466-894-367
; Sequence 367, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telesman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; TITLE OF INVENTION: AND THEIR USE AS MEDICINES
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 367
; LENGTH: 473
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 10, 431, 447, 467, 468, 473
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-367

Query Match 14.1%; Score 284.8; DB 18; Length 473;
Best Local Similarity 89.2%; Pred. No. 1.8e-64;
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGAGCGGAACTACTTGGACCAAGCGAGGACATGGAATGGGAAACCA 232
DB 72 CTTTACACGAGGAGCGGAGACCACTTGGACCAAGCGAGGACATGGAATGGGAGAGCG 131

QY 233 CAGTGGATGAATGACCAAAAGCAAGTATTCAATTTTGGATCTTTTGGTTAAAAATGTT 292
DB 132 CAGTGGATGACATGACCAAAAGCAAGTATTCAATTTTGGATCTTTTGGTTAAAAATGTT 191

QY 293 TATTTTGAAGTCTTAACACAAAGATATATTTCCCTGGTGATGTTAAATGTTTGGTTCGCAAC 352
DB 192 TGTTCGCTGTAATGTTCAACTAACACCAAGCTGAAAGAAATTTAA 516

QY 353 ATGAATGGGAAAAGCAAGCTGGCACTGGCATGTACTAAATGTTGAGGAAAGGACTTTA 412
DB 252 ATGAATGGGAAAAGCAAGCTGGCACTGTCTGTCTGATGTTGGAGGCAAGGACTTTA 311

QY 413 GTCAGCTCAAGGAAATGGTGAGAGCAAGCTGGCACTGGCATGTACTAAATGTTTACTGGAGCAGATGTTGG 472
DB 312 GTCAGCTCAAGGAAATGGTGAGAGCAAGCTGGCACTGGCATGTACTAAATGTTTACTGGAGCAGATGTTGG 371

QY 473 TAACAGCTGTAATGTGCAACTAACACCAAGCTGAAAGAAATTTAA 516

Db 372 TGACTGCTGTAATGTTCAACTAACACCGAGCTGAAAAA 415

RESULT 9

US-10-466-894-358
; Sequence 358, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; PRIOR FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 358
; LENGTH: 486
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 413, 428, 435, 448, 457, 462
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-358

Query Match 14.1%; Score 284.8; DB 18; Length 486;
Best Local Similarity 89.2%; Pred. No. 1.8e-64;
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGAGCGGAACTACTTTGGGACCAAGCGAGGACATGGAAATGGGAAACCA 232
Db 54 CTTTACACGAGGAGCGGAGACCACTTTGGGACCAAGCGAGGACATGGAAATGGGAGCG 113
QY 233 CAGTGGATGAATGACCAAAAGCAAGTATTCAATTTTGATCTTTTGGTTAAAAATGTT 292
Db 114 CAGTGGATGACATGACCAAAAGCAAGTATTATTTTGTGATCTTTTGGTTAAAGAGTGT 173
QY 293 TATTTGAAGTCTTAACACAAAGAAATATATTTCCCTGGTGTATTTGTTGTTGCAAC 352
Db 174 TGTTGAAGTCTTAACACAAAGAAATATATTTCCCTGGTGTATTTGTTGTTGCAAC 233
QY 353 ATGAATGGGAAAAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGGACTTTA 412
Db 234 ATGAATGGGAAAAGACCAAGGCTGGCACTGTCTGTGATTTGGAGGCAAGGACTTTA 293
QY 413 GTCAAGCTCAAGGAAATGGTGAGAGCAACTAAATGTTTACTTGAGCAGATGTTGG 472
Db 294 GTCAAGCTCAAGGAAATGGTGAGAGCAACTAAATGTTTACTTGAGTGTAGTGTGG 353
QY 473 TAACAGCTGTATGTGCAACTTAACACCAAGCTGAAAGAAATTTAA 516
Db 354 TGACTGCTGTATGTTCAACTTAACACCAAGCTGAAAGAAATTTAA 397

RESULT 10

US-10-466-894-428
; Sequence 428, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent

; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; PRIOR FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 428
; LENGTH: 491
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 12, 421, 441
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-428

Query Match 14.1%; Score 284.8; DB 18; Length 491;
Best Local Similarity 89.2%; Pred. No. 1.8e-64;
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGAGCGGAACTACTTTGGGACCAAGCGAGGACATGGAAATGGGAAACCA 232
Db 62 CTTTACACGAGGAGCGGAGACCACTTTGGGACCAAGCGAGGACATGGAAATGGGAGCG 121
QY 233 CAGTGGATGAATGACCAAAAGCAAGTATTCAATTTTGATCTTTTGGTTAAAAATGTT 292
Db 122 CAGTGGATGACATGACCAAAAGCAAGTATTATTTTGTGATCTTTTGGTTAAAGAGTGT 181
QY 293 TATTTGAAGTCTTAACACAAAGAAATATATTTCCCTGGTGTATTTGTTGTTGCAAC 352
Db 182 TGTTGAAGTCTTAACACAAAGAAATATATTTCCCTGGTGTATTTGTTGTTGCAAC 241
QY 353 ATGAATGGGAAAAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGGACTTTA 412
Db 242 ATGAATGGGAAAAGACCAAGGCTGGCACTGTCTGTGATTTGGAGGCAAGGACTTTA 301
QY 413 GTCAAGCTCAAGGAAATGGTGAGAGCAACTAAATGTTTACTTGAGCAGATGTTGG 472
Db 302 GTCAAGCTCAAGGAAATGGTGAGAGCAACTAAATGTTTACTTGAGTGTAGTGTGG 361
QY 473 TAACAGCTGTATGTGCAACTTAACACCAAGCTGAAAGAAATTTAA 516
Db 362 TGACTGCTGTATGTTCAACTTAACACCAAGCTGAAAGAAATTTAA 405

RESULT 11

US-10-466-894-438
; Sequence 438, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; PRIOR FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 438

us-10-069-056-8.rnpb

```
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 418_msc_1
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-438

Query Match      14.1%; Score 284.2; DB 18; Length 420;
Best Local Similarity 87.8%; Pred. No. 2.4e-64;
Matches 310; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 173 CTTTCAACGAGGCGGAACTACTTGGGACCAAGCGAGGACATGGAATGGGAACCA 232
Db 64 CTTTCAACGAGGCGGAGACCACTTGGGACCAAGCGAGGACATGGAATGGGAACCG 123
QY 233 CAGTGGATCAATGACCAAAAGCAAGTATTCATTTTGTGATTTTGTGTTAAATAATGT 292
Db 124 CAGTGGATGACATGACCAAAAGCAAGTATTTATTTTGTGATTTTGTGTTAAATAATGT 183
QY 293 TATTTGAAGTGCTTAACACAAAGATATATTTCTGGTGATGTTAATTTGGTTTGGCAAC 352
Db 184 TGTTCGAAGTGCTCAGCACAAAGCAATAGCTCCTAGTAATGTACTTGGTTTGTGCAGC 243
QY 353 ATGAATGGGAAAGACCAAGCTGCGACTGCCATGCTACTAATTTGAGGAAAGGACTTTA 412
Db 244 ATGAATGGGAAAGACCAAGCTGCGACTGCCATGCTACTAATTTGAGGAAAGGACTTTA 303
QY 413 GTCAAGCTCAAGGAAATGTTGAGAGGCACTAAATTTTACTGGACAGATGTTGG 472
Db 304 GTCAAGCTCAAGGAAATGTTGAGAGGCACTAAATTTTACTGGAGTAGATGTTGG 363
QY 473 TAAAGCTGCTTAATGTGCAACTAACACAGCTGAAAGATTAACCTAAAGAA 525
Db 364 TGACTGCTGCTTAATGTGCAACTAACACAGCTGAAAGATTAACCTAAAGAA 416

RESULT 12
US-10-466-894-388
; Sequence 388, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; TITLE OF INVENTION: AND THEIR USE AS MEDICINES
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; PRIOR FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 388
; LENGTH: 343
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-466-894-388

Query Match      14.1%; Score 283.8; DB 18; Length 343;
Best Local Similarity 89.2%; Pred. No. 2.7e-64;
Matches 306; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 174 TTTTCAACGAGGCGGAACTACTTGGGACCAAGCGAGGACATGGAATGGGAACCA 233
Db 1 TTTTCAACGAGGCGGAGACCACTTGGGACCAAGCGAGGACATGGAATGGGAACCGC 60
QY 234 AGTGGATGAATGACCAAAAGCAAGTATTCATTTTGTGATTTTGTGTTAAATAATGT 293
Db 61 AGTGGATGACATGACCAAAAGCAAGTATTTATTTTGTGATTTTGTGTTAAATAATGT 120
QY 294 ATTTGAAGTGCTTAACACAAAGATATATTTCTGGTGATGTTAATTTGGTTTGTGCAAC 353
Db 121 GTTTGAAGTGCTCAGCACAAAGCAATAGCTCCTAGTAATGTACTTGGTTTGTGCAAC 180
QY 354 TGAATGGGAAAGACCAAGCTGCGACTGCCATGCTACTAATTTGGAGGAAAGGACTTTAG 413
Db 181 TGAATGGGAAAGGACCAAGCTGCGACTGCCATGCTACTAATTTGGAGGAAAGGACTTTAG 240
QY 414 TCAAGCTCAAGGAAATGTTGAGAGGCACTAAATTTTACTGAGCAGATGTTGGT 473
Db 241 TCAAGCTCAAGGAAATGTTGAGAGGCACTAAATTTTACTGAGTAGATGTTGGT 300
QY 474 AACAGCTGTAATGTGCAACTAACACAGCTGAAAGATTAAC 516
```

```
QY 234 AGTGGATGAATGACCAAAAGCAAGTATTCATTTTGTGATTTTGTGTTAAATAATGT 293
Db 61 AGTGGATGACATGACCAAAAGCAAGTATTTATTTTGTGATTTTGTGTTAAATAATGT 120
QY 294 ATTTGAAGTGCTTAACACAAAGATATATTTCTGGTGATGTTAATTTGGTTTGTGCAAC 353
Db 121 GTTTGAAGTGCTCAGCACAAAGCAATAGCTCCTAGTAATGTACTTGGTTTGTGCAAC 180
QY 354 TGAATGGGAAAGACCAAGCTGCGACTGCCATGCTACTAATTTGGAGGAAAGGACTTTAG 413
Db 181 TGAATGGGAAAGGACCAAGCTGCGACTGCCATGCTACTAATTTGGAGGAAAGGACTTTAG 240
QY 414 TCAAGCTCAAGGAAATGTTGAGAGGCACTAAATTTTACTGAGCAGATGTTGGT 473
Db 241 TCAAGCTCAAGGAAATGTTGAGAGGCACTAAATTTTACTGAGTAGATGTTGGT 300
QY 474 AACAGCTGTAATGTGCAACTAACACAGCTGAAAGATTAAC 516
Db 301 GACTGCTGTAATGTGCAACTAACACAGCTGAAAGATTAAC 343

RESULT 13
US-10-466-894-390
; Sequence 390, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; TITLE OF INVENTION: AND THEIR USE AS MEDICINES
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; PRIOR FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 390
; LENGTH: 343
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-466-894-390

Query Match      14.1%; Score 283.8; DB 18; Length 343;
Best Local Similarity 89.2%; Pred. No. 2.7e-64;
Matches 306; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 174 TTTTCAACGAGGCGGAACTACTTGGGACCAAGCGAGGACATGGAATGGGAACCA 233
Db 1 TTTTCAACGAGGCGGAGACCACTTGGGACCAAGCGAGGACATGGAATGGGAACCGC 60
QY 234 AGTGGATGAATGACCAAAAGCAAGTATTCATTTTGTGATTTTGTGTTAAATAATGT 293
Db 61 AGTGGATGACATGACCAAAAGCAAGTATTTATTTTGTGATTTTGTGTTAAATAATGT 120
QY 294 ATTTGAAGTGCTTAACACAAAGATATATTTCTGGTGATGTTAATTTGGTTTGTGCAAC 353
Db 121 GTTTGAAGTGCTCAGCACAAAGCAATAGCTCCTAGTAATGTACTTGGTTTGTGCAAC 180
QY 354 TGAATGGGAAAGACCAAGCTGCGACTGCCATGCTACTAATTTGGAGGAAAGGACTTTAG 413
Db 181 TGAATGGGAAAGGACCAAGCTGCGACTGCCATGCTACTAATTTGGAGGAAAGGACTTTAG 240
QY 414 TCAAGCTCAAGGAAATGTTGAGAGGCACTAAATTTTACTGAGCAGATGTTGGT 473
Db 241 TCAAGCTCAAGGAAATGTTGAGAGGCACTAAATTTTACTGAGTAGATGTTGGT 300
QY 474 AACAGCTGTAATGTGCAACTAACACAGCTGAAAGATTAAC 516
```

Db 301 GACTGCTGTATGTTCAACTAACACGCTGAAAAA 343

RESULT 14
US-10-466-894-433
; Sequence 433, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; PRIOR FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 433
; LENGTH: 403
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 218
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-433

Query Match 14.1%; Score 283.8; DB 18; Length 403;
Best Local Similarity 89.0%; Pred. No. 2.9e-64;
Matches 306; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
Qy 173 CTTTACACGAGGAGCGGAACTACTTTGGACCAAGCGAGGACATGGAAATGGGAAACCA 232
Db 56 CTTTACACGAGGAGCGGAGACCACTTTGGACCAAGCGAGGACATGGAAATGGGAGCG 115
Qy 233 CAGTGGATGAATGACCAAAAGCAAGTATTCTTTTGGATCTTTTGGTTAAAAATGTT 292
Db 116 CAGTGGATGACATGACCAAAAGCAAGTATTCTTTTGGATCTTTTGGTTAAAGAGTGT 175
Qy 293 TATTTGAAGTGTCTTACACAAAGAAATATTTCTGCTGATGTTAATTTGGTTTGCAAC 352
Db 176 TGTTTGAAGTGTCTTACACAAAGAAATATTTCTGCTGATGTTAATTTGGTTTGCGCAGC 235
Qy 353 ATGAATGGGAAAGAACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGGACTTTA 412
Db 236 ATGAATGGGAAAGAACCAAGGCTGGCACTGTCTGTCTGTGATGGAGGCAAGGACTTTA 295
Qy 413 GTCAAGCTCAAGGAAATGTTGGAGAGGCACTAAATGTTTACTCGAGCAGATGTTGG 472
Db 296 GTCAAGCTCAAGGAAATGTTGGAGAGGCACTAAATGTTTACTCGAGTAGATGTTGG 355
Qy 473 TAACAGCTGTATGTTGCAACTAACACGCTGAAAGAAATTTAA 516
Db 356 TGACTGCTGTATGTTGCAACTAACACGCTGAAAGAAATTTAA 399

RESULT 15
US-10-466-894-369
; Sequence 369, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent

; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; PRIOR FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 369
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 448-475
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-369
Query Match 14.0%; Score 283.2; DB 18; Length 497;
Best Local Similarity 89.0%; Pred. No. 4.8e-64;
Matches 306; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
Qy 173 CTTTACACGAGGAGCGGAACTACTTTGGACCAAGCGAGGACATGGAAATGGGAAACCA 232
Db 68 CTTTACACGAGGAGCGGAGACCACTTTGGACCAAGCGAGGACATGGAAATGGGAGCG 127
Qy 233 CAGTGGATGAATGACCAAAAGCAAGTATTCTTTTGGATCTTTTGGTTAAAAATGTT 292
Db 128 CAGTGGATGACATGACCAAAAGCAAGTATTCTTTTGGATCTTTTGGTTAAAGAGTGT 187
Qy 293 TATTTGAAGTGTCTTACACAAAGAAATATTTCTGCTGATGTTAATTTGGTTTGCAAC 352
Db 188 TGTTTGAAGTGTCTTACACAAAGAAATATTTCTGCTGATGTTAATTTGGTTTGCGCAGC 247
Qy 353 ATGAATGGGAAAGAACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGGACTTTA 412
Db 248 ATGAATGGGAAAGAACCAAGGCTGGCACTGTCTGTCTGTGATGGAGGCAAGGACTTTA 307
Qy 413 GTCAAGCTCAAGGAAATGTTGGAGAGGCACTAAATGTTTACTCGAGCAGATGTTGG 472
Db 308 GTCAAGCTCAAGGAAATGTTGGAGAGGCACTAAATGTTTACTCGAGTAGATGTTGG 367
Qy 473 TAACAGCTGTATGTTGCAACTAACACGCTGAAAGAAATTTAA 516
Db 368 TGACTGCTGTATGTTGCAACTAACACGCTGAAAGAAATTTAA 411

Search completed: January 23, 2005, 14:35:11
Job time : 1074 secs

11

12

13

14

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2005, 23:46:58 ; Search time 6354.75 Seconds
(without alignments)
11577.453 Million cell updates/sec

Title: US-10-069-056-8

Perfect score: 2019

Sequence: 1 atggctggaatcttactc.....agccgtgaacttgactaa 2019

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	61.6	3.1	1101	9	CNS0039G
C 2	59.8	3.0	997	9	CNS005TE
C 3	50	2.5	1134	8	CC245800 CH261-4F2
C 4	49.8	2.5	1101	9	CNS0006J
C 5	49.8	2.5	1101	9	CNS0182P
C 6	49.6	2.5	1101	9	CNS017KX
C 7	49.4	2.4	886	9	CNS075DF
C 8	49.2	2.4	913	9	CG772152
C 9	49.2	2.4	930	9	CG770435
C 10	48.4	2.4	1063	9	CG770435 TC852.2_A
C 11	48.4	2.4	1241	9	AG448181
C 12	47.8	2.4	987	9	AL066537 Drosophila
C 13	47.2	2.3	1025	9	AL104216 Drosophila
C 14	47.2	2.3	1811	9	CG753732 P048-4-GO
C 15	46.6	2.3	933	9	AL059220 Drosophila
C 16	46.4	2.3	1080	9	CNS00EPP
C 17	46.2	2.3	1758	9	CL509408 SAIL_811
C 18	46	2.3	769	9	AL100817 Drosophila
C 19	45.8	2.3	963	9	CNS00A4L
C 20	45.8	2.3	1066	6	CD386086 AGENCOURT
C 21	45.8	2.3	1204	9	AL106628 Drosophila
C 22	45.6	2.3	1200	9	AL106647 Drosophila
C 23	45.4	2.2	975	6	CD050911 AGENCOURT
C 24	45.2	2.2	644	8	BH463121 BOHEK14TR

C 25	45.2	2.2	979	8	AZ538989
C 26	45	2.2	768	9	CNS01VSE
C 27	45	2.2	894	9	CNS01BEG
C 28	45	2.2	1184	8	B13117
C 29	44.8	2.2	368	2	AW767845
C 30	44.8	2.2	414	4	BJ640957
C 31	44.8	2.2	435	8	BZ665713
C 32	44.8	2.2	581	9	CL553259
C 33	44.6	2.2	298	4	B1815258
C 34	44.6	2.2	363	5	BQ451492
C 35	44.6	2.2	870	8	AZ693037
C 36	44.4	2.2	706	9	AG178839
C 37	44.4	2.2	878	9	CNS0187R
C 38	44.4	2.2	938	9	CNS006TJ
C 39	44.4	2.2	1201	5	BX416077
C 40	44.2	2.2	588	1	AJ773389
C 41	44.2	2.2	694	1	AJ778527
C 42	44.2	2.2	1500	9	AG334092
C 43	44	2.2	871	5	BU221434
C 44	44	2.2	922	9	CG766473
C 45	44	2.2	1101	9	CNS00240

ALIGNMENTS

RESULT 1
LOCUS CNS0039G/c
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL063921
VERSION AL063921.1 GI:4941778
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Ooeogawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone_lib="BACR08K10"
/note="end : TET3"

ORIGIN
Query Match 3.1%; Score 61.6; DB 9; Length 1101;
Best Local Similarity 16.2%; Pred. No. 0.00016;

AZ538989 ENTFFH71TF
AL186683 Tetradon
AL109126 Drosophila
B13117 T9K3-Sp6.2
AW767845 da80h01.x
BJ640957 BJ640957
BZ665713 KQ09816 D
CL553259 OB_Ba000
B1815258 PfESToaa1
BQ451492 PfESToaa0
AZ693037 ENTHP79TF
AG178839 Pan trogl
AL108993 Drosophila
AL065906 Drosophila
BX416077 BX416077
AJ773389 AJ773389
AJ778527 AJ778527
AG334092 Mus muscu
BU221434 603107334
CG766473 TCB41.1.H
AL061823 Drosophila

KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Genoscope.
REFERENCE Direct Submission
AUTHORS Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.
FEATURES Location/Qualifiers
source 1. .1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN37F10"
/library="DrosBAC"
/plasmid="pBelobAC11"
/notes="end : Sp6"
ORIGIN
Query Match 2.5%; Score 49.6; DB 9; Length 1101;
Best Local Similarity 21.8%; Pred. No. 0.21;
Matches 95; Conservative 143; Mismatches 198; Indels 0; Gaps 0;
QY 260 TATTCTCTGATGTTTAAATGTTTGTGCAATGAATGGGAAAGACCAAGCGTGC 379
Db 1077 TWTADUKTWRTDKDGGGGGKGGKGGDRKADTKRDWRGAATRRARWATAATTATA 1018
QY 320 TATTCTCTGATGTTTAAATGTTTGTGCAATGAATGGGAAAGACCAAGCGTGC 379
Db 1017 DDTWKRTTGTTRTKWKWWAREAKWRDWRDWDKWDTKWKDDRTWKRWATRGDGR 958
QY 380 ACTGCATGTAATTTGAGGAAAGACCTTTAGTCAAGCTCAAGGAAATGTTGAGAA 439
Db 957 AGRKERDRGRGRRRRRTTRKWRWRRAAWTAWTTTMMGGTTKGAWTTKTRKKG 898
QY 440 GGCACCTAAATGTTTACTGCGAGAGATGTTGTCAAGCTGTAATGTGCACTAACAC 499
Db 897 GAAAAADGRTDWRWDKRAAAAGKDGKKGKARTWGAAAAAAWDTKKTATKRRKRT 838
QY 500 CAGCTGAAAGAAATAAAGTAAAGAAATAGCAGACAAATGATGGGTACTCTACTTA 559
Db 837 KTGRAWAAAGRTTWDAAATWTKRAAAAGRAAAWAKDRGTGKGGKGGTDDGKG 778
QY 560 CTTATAGCAATAGCAAAACCAAAAAAGACTATACCAAGTGTGTTCTTTTGGAAACATGA 619
Db 777 TWKTRTTTWTATAAARRARAAAKGTTKTDGRRAAAWDAWTTTTRKGAATRDGAA 718
QY 620 TTGCTTACTATTTTAACTAAAGAAATAGACTAGTCCACCAAGACGGAGGCT 679
Db 717 AAAATTTGGAAAAATCGKTTDAGAAAAATAAWRRTTWTDWAAAAAAGAAWAKAAAWD 658
QY 680 ATTTCTTTAGCAGTGA 695
Db 657 AAADAKKGGRRAKRA 642
RESULT 7
CNS075DF 886 bp DNA linear GSS 07-JUL-2001
LOCUS clone BA0AB034H01 of library BA0AB from strain CLIB 210 of
DEFINITION Kluyveromyces lactis, genomic survey sequence.

AL429961 GI:12213155
VERSION GSS.
SOURCE Kluyveromyces lactis
ORGANISM Kluyveromyces lactis
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
1 (bases 1 to 886)
REFERENCE Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G., Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S., DeMontigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B., Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S., Saurin,W., Tekai,F., Toffano-Nioche,C., Wesolowski-Louvel,M., Wincker,P. and Weissenbach,J.
COMMENT Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies
FEB5 Lett. 487 (1), 3-12 (2000)
20584711
11152876
2 (bases 1 to 886)
REFERENCE Bolotin-Fukuhara,M., Toffano-Nioche,C., Artiguenave,F., Duchateau-Nguyen,G., Lemaire,M., Marmeisse,R., Montrocher,R., Robert,C., Termier,M., Wincker,P. and Wesolowski-Louvel,M.
COMMENT Genomic exploration of the hemiascomycetous yeasts: 11.
Kluyveromyces lactis
FEB5 Lett. 487 (1), 66-70 (2000)
20584721
11152886
3 (bases 1 to 886)
REFERENCE Genoscope.
DIRECT SUBMISSION
TITLE Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxi, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.
FEATURES Location/Qualifiers
source 1. .886
/organism="Kluyveromyces lactis"
/mol_type="genomic DNA"
/strain="CLIB 210"
/variety="lactis"
/db_xref="taxon:28985"
/clone="BA0AB034H01"
/clone_lib="BA0AB"
complement(<619..>886)
misc_feature /note="similar to Saccharomyces cerevisiae ORF YNR001c [cit1 ; citrate (st)-synthase, mitochondrial]"
/evidence=not_experimental
ORIGIN
Query Match 2.4%; Score 49.4; DB 9; Length 886;
Best Local Similarity 35.5%; Pred. No. 0.22;
Matches 196; Conservative 64; Mismatches 291; Indels 1; Gaps 1;
QY 491 AACTAACACAGCTGAAGAAATTAAGTAAGAAATAGCAGAAACAATGAGTGGTTA 550
Db 34 AAAAAACGAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 93
QY 551 CTCTACTTACTTATAAGCATAGCAAAACCAAAAAAGACTATACCAAGTGTGTTCTTTG 610
Db 94 AAAAAAATAACAAAAACAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 153
QY 611 GAAACATGATTCCTTACTATTTTAACTAAAGAAATAAGCACTAGTCCACCAAGAG 670

```
Db 154 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 213
Qy 671 ACGGAGCTATTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAAAGAGGGCG 730
Db 214 ARAAARTRTGAGAPAAATRWRTGRTGTGTAWRWWRWRTTATRWATRAAGWR 273
Qy 731 AGCGCATCTAGTGAGCAAACTATACACTGTATGACATGCGGCGAGAAACGGTTGAAACCA 790
Db 274 AAWRAAAATGWRAGGAGGAGGAAGARATATATAAARAAGARAARAAATWAAAWAR 333
Qy 791 CA-GTAACCACTGCGCAGAAACTAAGCCGCGGAGAAATCAAACTAAAGAAAGTTTCT 849
Db 334 TAGGTWAAAAAGTATWARWAWAWAAGAAAAAARAAAAAAMAAAAAAMAAAAAAMGGG 393
Qy 850 ATTAAACCTACACTTAAAGAGCTGGTGCAATAAAGAGTAACCTCACAGGAGCTGGATG 909
Db 394 CRGMAAAMAAAMAAAMAAACGTAGARGAAAMAAAMAAAMAAAMAAAMAAAMAAAGAR 453
Qy 910 ATGATGCAGCCAGACAGTTTACATTGAAATGATGGCTCAACACGAGTGAGAAAAACCTGCTG 969
Db 454 AARGACTTWRGNWAWAMGAATAGAAAAMCAACGMAAAAMGAARAAAMAAAMAAAMAAAR 513
Qy 970 AAAAAACGCTAGAGATTTGTACACTTAATCTTAGCCAGAACCAAAACACATTTGACTTA 1029
Db 514 AAMCFEAAAACMAAAAMAAACAMWRAAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAACCR 573
Qy 1030 ATTTTAGAAAAA 1041
Db 574 AAGTAAAMAAAA 585

RESULT 8
CG772152 913 bp DNA linear GSS 29-OCT-2003
LOCUS TcB48.2_H01_T7 Tribolium BAC library Tribolium castaneum genomic,
DEFINITION genomic survey sequence.
ACCESSION CG772152 GI:38025852
VERSION CG772152.1
KEYWORDS Tribolium castaneum (red flour beetle)
SOURCE Tribolium castaneum
ORGANISM Tribolium castaneum
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
Tenebrionidae; Tribolium.
REFERENCE 1 (bases 1 to 913)
AUTHORS Savard, J. and Tautz, D.
TITLE Tribolium castaneum BAC-ends sequencing project
JOURNAL Unpublished (2003)
COMMENT Abteilung fur Evolutionsgenetik, AG Tautz
Institut fur Genetik, Universitat zu Koln
Weyertal 121, 50931 Koln, Germany
Tel: 49 221 470 6911
Fax: 49 221 470 5975
Email: savard@uni-koeln.de
Class: BAC ends.

FEATURES
source
location/Qualifiers
1..913
/organism="Tribolium castaneum"
/mol_type="genomic DNA"
/strain="GA-2"
/db_xref="taxon:7070"
/clone_lib="Tribolium BAC library"
/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
Library constructed by Exelixis Inc."

ORIGIN
Query Match 2.4%; Score 49.2; DB 9; Length 913;
Best Local Similarity 58.0%; Pred. No. 0.25; Indels 0; Gaps 0;
Matches 87; Conservative 0; Mismatches 63;

Qy 23 ATGAAGTTTGGAGCAACCACTGGTTAAAGAAAAAGTAACCAAGGAGTGTCTCAT 82
Db 288 ATGAATTTTACACAGACCAATTAATACATCTTTAAATGATTTGGAGTATTCTGAA 347
Qy 83 TTGTTTTTAAAAATGAAATGTTCAACTGAATGGAAGAAATATCGGATGGAATAGTTACA 142
Db 348 CTATTTTATAACAAAAATCTTCGAAAGACAGAAAAATAAATATCGACTACTAAA 407
Qy 143 AAAAGAGCTGCAGGAGGACGAGCTGAAAT 172
Db 408 AAATAACTATCGACAACCTCTGTACTAAAAT 437

RESULT 10
CNS07A2Y 1063 bp DNA linear GSS 08-JUL-2001
LOCUS T3 end of clone BC0AA001B01 of library BC0AA from strain CBS 767 of
DEFINITION Debaryomyces hansenii, genomic survey sequence.
ACCESSION AL436064
VERSION AL436064.1 GI:12219477
KEYWORDS GSS.
SOURCE Debaryomyces hansenii (anamorph: Candida famata)
```

```
Db 296 ATGAATTTTACACAGACCAATTAATACATCTTTAAATGATTTGGAGTATTCTGAA 355
Qy 83 TTGTTTTTAAAAATGAAATGTTCAACTGAATGGAAGAAATATCGGATGGAATAGTTACA 142
Db 356 CTATTTTATAACAAAAATCTTCGAAAGAAACAGAAAAATAAATATCGACTACTAAA 415
Qy 143 AAAAGAGCTGCAGGAGGACGAGCTGAAAT 172
Db 416 AAATAACTATCGACAACCTCTGTACTAAAAT 445

RESULT 9
CG770435 930 bp DNA linear GSS 29-OCT-2003
LOCUS TcB52.2_A06_Sp6 Tribolium BAC library Tribolium castaneum genomic,
DEFINITION genomic survey sequence.
ACCESSION CG770435
VERSION CG770435.1 GI:38023614
KEYWORDS Tribolium castaneum (red flour beetle)
SOURCE Tribolium castaneum
ORGANISM Tribolium castaneum
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
Tenebrionidae; Tribolium.
REFERENCE 1 (bases 1 to 930)
AUTHORS Savard, J. and Tautz, D.
TITLE Tribolium castaneum BAC-ends sequencing project
JOURNAL Unpublished (2003)
COMMENT Abteilung fur Evolutionsgenetik, AG Tautz
Institut fur Genetik, Universitat zu Koln
Weyertal 121, 50931 Koln, Germany
Tel: 49 221 470 6911
Fax: 49 221 470 5975
Email: savard@uni-koeln.de
Class: BAC ends.

FEATURES
source
location/Qualifiers
1..930
/organism="Tribolium castaneum"
/mol_type="genomic DNA"
/strain="GA-2"
/db_xref="taxon:7070"
/clone_lib="Tribolium BAC library"
/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
Library constructed by Exelixis Inc."

ORIGIN
Query Match 2.4%; Score 49.2; DB 9; Length 930;
Best Local Similarity 58.0%; Pred. No. 0.25; Indels 0; Gaps 0;
Matches 87; Conservative 0; Mismatches 63;

Qy 23 ATGAAGTTTGGAGCAACCACTGGTTAAAGAAAAAGTAACCAAGGAGTGTCTCAT 82
Db 288 ATGAATTTTACACAGACCAATTAATACATCTTTAAATGATTTGGAGTATTCTGAA 347
Qy 83 TTGTTTTTAAAAATGAAATGTTCAACTGAATGGAAGAAATATCGGATGGAATAGTTACA 142
Db 348 CTATTTTATAACAAAAATCTTCGAAAGACAGAAAAATAAATATCGACTACTAAA 407
Qy 143 AAAAGAGCTGCAGGAGGACGAGCTGAAAT 172
Db 408 AAATAACTATCGACAACCTCTGTACTAAAAT 437

RESULT 10
CNS07A2Y 1063 bp DNA linear GSS 08-JUL-2001
LOCUS T3 end of clone BC0AA001B01 of library BC0AA from strain CBS 767 of
DEFINITION Debaryomyces hansenii, genomic survey sequence.
ACCESSION AL436064
VERSION AL436064.1 GI:12219477
KEYWORDS GSS.
SOURCE Debaryomyces hansenii (anamorph: Candida famata)
```

ORGANISM	Debaryomyces hansenii
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Debaryomyces.	
REFERENCE	1 (bases 1 to 1063)
AUTHORS	Souciet J.L., Aigle M., Artiguenave F., Blandin G., Bolotin-Fukuhara M., Bon E., Brottier P., Casaregola S., de-Montigny J., Dujon B., Durrens P., Lepingle A., Llorente B., Malpertuy A., Neuveglise C., Ozier-Kalogeropoulos O., Pottier S., Saurin W., Tekala F., Toffano-Nioche C., Wesolowski-Louvel M., Winkler P. and Weissenbach J.
TITLE	Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies
JOURNAL	FEBS Lett. 487 (1), 3-12 (2000)
MEDLINE	20584711
PUBMED	11152876
REFERENCE	2 (bases 1 to 1063)
AUTHORS	Lepingle A., Casaregola S., Neuveglise C., Bon E., Nguyen H., Artiguenave F., Winkler P. and Gaillardin C.
TITLE	Genomic exploration of the hemiascomycetous yeasts: 14. Debaryomyces hansenii var. hansenii
JOURNAL	FEBS Lett. 487 (1), 82-86 (2000)
MEDLINE	20584724
PUBMED	11152889
REFERENCE	3 (bases 1 to 1063)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (08-SP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT	This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.
FEATURES	Location/Qualifiers
source	1..1063
	/organism="Debaryomyces hansenii"
	/mol_type="genomic DNA"
	/strain="CBS 767"
	/variety="hansenii"
	/db_xref="taxon:4959"
	/clone="BC0A001B01"
	/clone_lib="BC0A"
	/note="end : T3"
ORIGIN	
Query Match	2.4%; Score 48.4; DB 9; Length 1063;
Best Local Similarity	38.2%; Pred.No. 0.42; Indels 0; Gaps 0;
Matches	108; Conservative
Qy	369 CCAAGGCTGGCACTGCCATCTACTAATTTGGAGGAAGGACTTTAGTCAAGCTCAAGGAA 428
Db	661 CCCGSSSSSSACCGCTTGGNTTCTGTGSGTGSITTTTTTTTWWWKWWATWYAT 720
Qy	429 ATGGTGGAGGAAGCACTAAATGTTTACTGAGCAGATGGTTGGTACACGCCTGTATGT 488
Db	721 TAAATGGATATTTTTRDDGGTTTGTATTAAGGGGRRRTATRTDTKTCTTTTGGGGGRRRA 780
Qy	489 GCAACTAACACCGAGCTGAAAGAAATTAACCTAAGAGAATAGCAGACAACTAGTGGGT 548
Db	781 RRARRRRGAARRGGRRGGGKGHAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 840
Qy	549 TACTCTACTTACTTATTAAGCATAGCAAAACCAAAAAAGACTATACCAAGTGTGTCTTT 608
Db	841 AAATTTTTTAAAAAAGAAAAAATAAATAAAGAAAAAATTTTWTATTTTATATTTT 900
Qy	609 TGGAAACATGATTGCTTACTATTTTTTAACTAAAGAAAAA 651

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:18:38 ; Search time 99 Seconds
(without alignments)
2435.011 Million cell updates/sec

Title: US-10-069-056-10

Perfect score: 3637

Sequence: 1 MAGNAYSDEVLGATNLKKE.....RACFGAEPLKDFSEPLNLD 672

Scoring table: BLOSOM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3627	99.7	672	4	AAY72706 Parvoviru
2	3623	99.6	672	4	AAY72702 Parvoviru
3	3620	99.5	672	4	AAY72704 Parvoviru
4	3618	99.5	672	4	AAY72708 Parvoviru
5	3618	99.5	672	4	AAY72710 Parvoviru
6	2371	65.2	690	1	AAP40306 Sequence
7	458.5	12.6	550	3	AAY58164 Adeno ass
8	458.5	12.6	550	5	AAU11409
9	458.5	12.6	610	3	AAY58159 Adeno ass
10	458.5	12.6	610	4	AAY97720 Rep prote
11	458.5	12.6	610	5	AAU11404 Adeno-ass
12	458.5	12.6	610	5	AAE22887 Adeno-ass
13	458.5	12.6	610	5	AAE28644 Adeno-ass
14	458.5	12.6	610	5	AAE26940 Adeno ass
15	458.5	12.6	610	6	ABU64865 Rep prote
16	458.5	12.6	610	6	ABU64760 Adeno ass
17	458.5	12.6	610	6	ABR43398 Adeno-ass
18	458.5	12.6	610	7	ADI40280 Adeno-ass
19	458.5	12.6	610	8	ADH58895 Predeterm
20	455.5	12.5	626	2	AAE28648 Barbary d
21	455.5	12.5	626	4	AAAY97724 NSI prote
22	455.5	12.5	626	5	AAE22891 Muscovy d
23	455.5	12.5	626	5	AAE28648 Muscovy d
24	455.5	12.5	626	5	AAE26944 Muscovy d
25	455.5	12.5	626	6	ABU64869 NSI prote

26	455.5	12.5	626	6	ABU64764	Abu64764 Parvoviru
27	455.5	12.5	626	6	ABR43402	AbR43402 Muscovy d
28	455.5	12.5	626	7	ADI40288	Adi40288 Muscovy d
29	455	12.5	627	4	AAAY97722	Aay97722 Rep prote
30	455	12.5	627	5	AAE22889	Aae22889 Barbare
31	455	12.5	627	5	AAE28646	Aae28646 Barbare
32	455	12.5	627	5	AAE26942	Aae26942 Barbare
33	455	12.5	627	6	ABU64867	Abu64867 Rep prote
34	455	12.5	627	6	ABU64762	Abu64762 Parvoviru
35	455	12.5	627	6	ABR43400	AbR43400 Barbare
36	455	12.5	627	7	ADI40284	Adi40284 Barbare
37	445.5	12.2	625	7	ABR62760	AbR62760 Adeno ass
38	437.5	12.0	536	2	AAW46311	Aaw46311 AAV4 Rep
39	437.5	12.0	536	6	ABG73938	AbG73938 Adeno-ass
40	437.5	12.0	623	2	AAW46307	Aaw46307 AAV4 Rep
41	437.5	12.0	623	2	AAW46312	Aaw46312 AAV4 Rep
42	437.5	12.0	623	4	AAAY97712	Aay97712 Rep prote
43	437.5	12.0	623	5	AAE28636	Aae28636 Adeno-ass
44	437.5	12.0	623	5	AAE26932	Aae26932 Adeno ass
45	437.5	12.0	623	6	ABU64857	Abu64857 Rep prote

ALIGNMENTS

RESULT 1

AAAY72706

ID AAY72706 standard; protein; 672 AA.

XX AAY72706;

AC AAY72706;

XX

DT 06-AUG-2003 (revised)

DT 31-MAY-2001 (first entry)

XX

DE Parvovirus non-structure protein 1 (NS1) variant (T363A).

XX

KW NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;

KW tumoural disease; gene therapy; mutant; mutein; variant.

XX

OS Parvovirus.

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Misc-difference 363

FT /note= "Wild type Thr substituted with Ala"

XX

PN EP1077260-A1.

XX

PD 21-FEB-2001.

XX

PF 13-AUG-1999; 99EP-00115161.

XX

PR 13-AUG-1999; 99EP-00115161.

XX

PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX

PI Nuesch J, Rommelaere J;

XX

DR WPI: 2001-212717/22.

DR

DR N-PSDB; AAD02801.

XX

PT Novel parvovirus non-structure protein variant, useful for treating

PT tumoral diseases, has a shifted equilibrium between DNA replication and

PT transcription activities, and cytotoxic activity.

XX

PS Claim 6; Page 19-21; 41pp; English.

XX

CC The present sequence is parvovirus non-structure protein 1 (NS1) variant

CC (T363A). The invention relates to the variants of the parvovirus non-

CC structure protein (NS1) having a shifted equilibrium between the DNA

CC replication and transcription activities, and the cytotoxicity activity.

CC These variants are useful as vectors for treating tumoural diseases. The

CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-

CC 2003 to correct OS field.)

XX
SQ Sequence 672 AA;

Query Match 99.7%; Score 3627; DB 4; Length 672;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 671; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MAGNAYSDEVLGATNLKESNQEVFSFVKNNVQLNGKOIGWNSYKKELQEDELKSQ 60
Dd |||||
Qy 1 MAGNAYSDEVLGATNLKESNQEVFSFVKNNVQLNGKOIGWNSYKKELQEDELKSQ 60
Dd |||||

Qy 61 RGAETTWDSQSDMEWETTVDENTKKQVFIFDSLKKCLFEVLNTKNIPFGDVNFVQHEW 120
Dd |||||

Qy 61. RGAETTWDSQSDMEWETTVDENTKKQVFIFDSLKKCLFEVLNTKNIPFGDVNFVQHEW 120
Dd |||||

Qy 121 GKDGWHCHLVIGGKDFSOAQGWRRQLNVYSRWLVTACNVLTTPAERIKLEIAEDN 180
Dd |||||

Qy 121 GKDGWHCHLVIGGKDFSOAQGWRRQLNVYSRWLVTACNVLTTPAERIKLEIAEDN 180
Dd |||||

Qy 181 EWTLLTYKHQTKDYTCVLFGNMAYFYFTKKISTSPRGGYFLSSDSGWKNFL 240
Dd |||||

Qy 181 EWTLLTYKHQTKDYTCVLFGNMAYFYFTKKISTSPRGGYFLSSDSGWKNFL 240
Dd |||||

Qy 241 KEGEHLVSKLYTDDMRPETVTTTAQETKRGRIOQTKEVSIKITTLKELVHKRVTSPE 300
Dd |||||

Qy 241 KEGEHLVSKLYTDDMRPETVTTTAQETKRGRIOQTKEVSIKITTLKELVHKRVTSPE 300
Dd |||||

Qy 301 DWMMPQDSYIENMAQPGENLLKNTLCITLTARTKTAFDLILEKAETSCLTNFSLPD 360
Dd |||||

Qy 301 DWMMPQDSYIENMAQPGENLLKNTLCITLTARTKTAFDLILEKAETSCLTNFSLPD 360
Dd |||||

Qy 361 TRACRIFAFHGWNYKVCHAICCVLNROGKRNVTLPFGPASTGKSIIQAIAOAVGVNG 420
Dd |||||

Qy 361 TRACRIFAFHGWNYKVCHAICCVLNROGKRNVTLPFGPASTGKSIIQAIAOAVGVNG 420
Dd |||||

Qy 421 CYNAAVNPFNDCTNKLIWVEEAGNFQQVNOFKAI CSQTIRIDQKGSGKIIEPTP 480
Dd |||||

Qy 421 CYNAAVNPFNDCTNKLIWVEEAGNFQQVNOFKAI CSQTIRIDQKGSGKIIEPTP 480
Dd |||||

Qy 481 VIMTTNENITTVRIGCEERPEHTQPIRDRLMLNIHILTHLPDGLVDKNEWPMI CAMLVK 540
Dd |||||

Qy 481 VIMTTNENITTVRIGCEERPEHTQPIRDRLMLNIHILTHLPDGLVDKNEWPMI CAMLVK 540
Dd |||||

Qy 541 NGYSTMASYCAKWGPDPWSNWAEPKVPPTINILGSARSPTTPKSTPLSQNYALTPL 600
Dd |||||

Qy 541 NGYSTMASYCAKWGPDPWSNWAEPKVPPTINILGSARSPTTPKSTPLSQNYALTPL 600
Dd |||||

Qy 601 ASDLEDALPEWSTPNTPVAGTAETQNTGEAGSKACQDGLSPTWSEIEDLRACFGAEP 660
Dd |||||

Qy 601 ASDLEDALPEWSTPNTPVAGTAETQNTGEAGSKACQDGLSPTWSEIEDLRACFGAEP 660
Dd |||||

Qy 661 LKXDFSEPLNLD 672
Dd |||||

Qy 661 LKXDFSEPLNLD 672
Dd |||||

RESULT 2
AA72702
ID AAY72702 standard; protein; 672 AA.
XX
XX
XX AAY72702;
XX AC
XX CC
DT 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
XX Parvovirus wild-type non-structure protein 1 (NS1).
XX DE
XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoural disease; gene therapy.
XX
XX Parvovirus.
XX OS
XX

Db 541 NGYQSTMASYCAKWKVPDSENWAEKVPPTPINLLGSARSPTTPKSTPLSQNYALTPL 600
 QY 601 ASDLEDLALPEWSTNTPVAGTAETONTGEAGSKACQDQGLSPTWSEIEDLRACFGARP 660
 Db 601 ASDLEDLALPEWSTNTPVAGTAETONTGEAGSKACQDQGLSPTWSEIEDLRACFGARP 660
 QY 661 LKKDFSEPLNLD 672
 Db 661 LKKDFSEPLNLD 672

RESULT 3
 AAY72704
 ID AAY72704 standard; protein; 672 AA.
 AC AAY72704;
 DT 06-AUG-2003 (revised)
 DT 31-MAY-2001 (first entry)
 XX Parvovirus non-structure protein 1 (NS1) variant (S283A).
 DE NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
 KW tumoral disease; gene therapy; mutant; mutein; variant.
 XX Parvovirus.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc-difference 283 /note= "wild type Ser substituted with Ala"
 FT EP1077260-Al.
 PN 21-FEB-2001.
 PD 13-AUG-1999; 99EP-00115161.
 XX 13-AUG-1999; 99EP-00115161.
 PR (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 PA Nueesch J, Rommelaere J;
 PI WPI; 2001-212717/22.
 DR N-PSDB; AAD02799.
 XX Novel parvovirus non-structure protein variant, useful for treating
 PT tumoral diseases, has a shifted equilibrium between DNA replication and
 PT transcription activities, and cytotoxic activity.
 XX Claim 6; Page 14-16; 41pp; English.
 CC The present sequence is parvovirus non-structure protein 1 (NS1) variant
 CC (S283A). The invention relates to the variants of the parvovirus non-
 CC structure protein (NS1) having a shifted equilibrium between the DNA
 CC replication and transcription activities, and the cytotoxicity activity.
 CC These variants are useful as toxins for treating tumoral diseases. The
 CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
 CC 2003 to correct OS field.)
 XX Sequence 672 AA;

Query Match 99.5%; Score 3620; DB 4; Length 672;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 669; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVGATNWLKESKQEVFSFVFKENVQLNGKDIGNWSYKKEQLQDELKSLQ 60
 Db 1 MAGNAYSDEVGATNWLKESKQEVFSFVFKENVQLNGKDIGNWSYKKEQLQDELKSLQ 60
 QY 61 RGAETTTWOSEDMEWETTVDENTKKQVFI FDSLVKKCLFEVLNTKNI FPGDVNWFVQHEW 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 61 RGAETTTWOSEDMEWETTVDENTKKQVFI FDSLVKKCLFEVLNTKNI FPGDVNWFVQHEW 120
 QY 121 GKQDGHCHVLIGGKDFSOAQGWRRQLUNVYWSRLVLTACNVQLTPAERIKLREIAEDN 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 121 GKQDGHCHVLIGGKDFSOAQGWRRQLUNVYWSRLVLTACNVQLTPAERIKLREIAEDN 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 181 EWTLLTYKHQTKKDYTKCVLFGNMIAYYFTLTKKISTSPRDGGYFLSSDSGWTNPL 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 181 EWTLLTYKHQTKKDYTKCVLFGNMIAYYFTLTKKISTSPRDGGYFLSSDSGWTNPL 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 241 KEGERHLSKLYTDDMRPETVTTTAQETKRGRIQTKEYSIKTTLKLHVHKTSP 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 241 KEGERHLSKLYTDDMRPETVTTTAQETKRGRIQTKEVAIKTTLLKELVHKRTSP 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 301 DMMWQPDYSIEMMAQPGGENLLKNTLEICTLTLTARTKTAFDLILEKAETSKLTNFSLPD 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 301 DMMWQPDYSIEMMAQPGGENLLKNTLEICTLTLTARTKTAFDLILEKAETSKLTNFSLPD 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 361 TRACRIFAPHGWNVYKVAICCVLNROGGKENTVLFHGPASTGKSIIAQAIAQAVGNVG 420
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 361 TRTCRIFAPHGWNVYKVAICCVLNROGGKENTVLFHGPASTGKSIIAQAIAQAVGNVG 420
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 421 CYNAANVAPFPNDCTNKNLIWVEEAGNFQGVNQPKAICSGQTIRIDQKSGSKQIEPTP 480
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 421 CYNAANVAPFPNDCTNKNLIWVEEAGNFQGVNQPKAICSGQTIRIDQKSGSKQIEPTP 480
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNWPMICAWLVK 540
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNWPMICAWLVK 540
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 541 NGYQSTMASYCAKWKVPDSENWAEKVPPTPINLLGSARSPTTPKSTPLSQNYALTPL 600
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 541 NGYQSTMASYCAKWKVPDSENWAEKVPPTPINLLGSARSPTTPKSTPLSQNYALTPL 600
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 601 ASDLEDLALPEWSTNTPVAGTAETONTGEAGSKACQDQGLSPTWSEIEDLRACFGARP 660
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 601 ASDLEDLALPEWSTNTPVAGTAETONTGEAGSKACQDQGLSPTWSEIEDLRACFGARP 660
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 661 LKKDFSEPLNLD 672
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 661 LKKDFSEPLNLD 672

RESULT 4
 AAY72708
 ID AAY72708 standard; protein; 672 AA.
 AC AAY72708;
 XX 06-AUG-2003 (revised)
 DT 31-MAY-2001 (first entry)
 XX Parvovirus non-structure protein 1 (NS1) variant (T394A).
 DE NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
 KW tumoral disease; gene therapy; mutant; mutein; variant.
 XX Parvovirus.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc-difference 394 /note= "Wild type Thr substituted with Ala"
 FT EP1077260-Al.
 PN 21-FEB-2001.
 PD 13-AUG-1999; 99EP-00115161.
 XX 13-AUG-1999; 99EP-00115161.
 PR 13-AUG-1999; 99EP-00115161.
 XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 PA

```
..XX
PI Nueesch J, Rommelaere J;
XX WPI: 2001-212717/22.
DR N-PSDB; AAD02805.
XX
PT Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX
PS Claim 6; Page 25-27; 41pp; English.
XX
CC The present sequence is parvovirus non-structure protein 1 (NS1) variant
CC (T394A). The invention relates to the variants of the parvovirus non-
CC structure protein (NS1) having a shifted equilibrium between the DNA
CC replication and transcription activities, and the cytotoxicity activity.
CC These variants are useful as toxins for treating tumoral diseases. The
CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX
SQ Sequence 672 AA;

Query Match 99.5%; Score 3618; DB 4; Length 672;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 669; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVLGATNWLKEKSNQEVFSFVKNNVQLNGKDIGNWSYKKELODELKSLQ 60
DB 1 MAGNAYSDEVLGATNWLKEKSNQEVFSFVKNNVQLNGKDIGNWSYKKELODELKSLQ 60
QY 61 RGAETTWQSEDMEWETTVDMTKQVIFPDSLVLKVKCLFEVLNTKNIFFGDVNVFVQHEW 120
DB 61 RGAETTWQSEDMEWETTVDMTKQVIFPDSLVLKVKCLFEVLNTKNIFFGDVNVFVQHEW 120
QY 121 GKQDQGHCHVLIGKDFSOAQGWRRQLNYSRWLVTAACNVQLTPAERIKLREIAEDN 180
DB 121 GKQDQGHCHVLIGKDFSOAQGWRRQLNYSRWLVTAACNVQLTPAERIKLREIAEDN 180
QY 181 EWTLLTYKHQTKKDYTKCVLFGNMIAYYFLTKKISTSPRDGGYFLSSDSGKWTNPL 240
DB 181 EWTLLTYKHQTKKDYTKCVLFGNMIAYYFLTKKISTSPRDGGYFLSSDSGKWTNPL 240
QY 241 KEGERHLVSKLYTDDMRPEVETVTTAQTETKRGRIQTKKEVSIKTTLLKELVHKRVTSPE 300
DB 241 KEGERHLVSKLYTDDMRPEVETVTTAQTETKRGRIQTKKEVSIKTTLLKELVHKRVTSPE 300
QY 301 DNMMPDSDSYIENMAOPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLEP 360
DB 301 DNMMPDSDSYIENMAOPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLEP 360
QY 361 TRACRIFAHGWNVYVCHAI CCVLNRQGGKRNVTFLFHPGASTGKSIIAQAIAQAVGNVG 420
DB 361 TRCIRIFAHGWNVYVCHAI CCVLNRQGGKRNAVL FHPGASTGKSIIAQAIAQAVGNVG 420
QY 421 CYNAAVNVPFNDCTKNLIWEEAGNFGQVQVNFKAICSGQTIRIDQKGKSKQIEPTP 480
DB 421 CYNAAVNVPFNDCTKNLIWEEAGNFGQVQVNFKAICSGQTIRIDQKGKSKQIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNEMPMI CAWLK 540
DB 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNEMPMI CAWLK 540
QY 541 NGYQSTMASYCAKWKVPDWSNWAEPKVPPTPINLIGSARSPFTTPKSTPLSQNTALTP 600
DB 541 NGYQSTMASYCAKWKVPDWSNWAEPKVPPTPINLIGSARSPFTTPKSTPLSQNTALTP 600
QY 601 ASDLELALPWSPTNPVAGTAETONTGEAGSKACQDQGLSPTWSEI EEDLRACFGAEP 660
DB 601 ASDLELALPWSPTNPVAGTAETONTGEAGSKACQDQGLSPTWSEI EEDLRACFGAEP 660
QY 661 LKXDFSEPLNLD 672
DB 661 LKXDFSEPLNLD 672
```

```
RESULT 5
AA72710
ID AAY72710 standard; protein; 672 AA.
XX
AC AAY72710;
XX
DT 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
DE Parvovirus non-structure protein 1 (NS1) variant (T463A).
XX
KW NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoral disease; gene therapy; mutant; mutein; variant.
XX
OS Parvovirus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 463 /note= "Wild type Thr substituted with Ala"
FT
XX EP1077260-A1.
PN
XX 21-FEB-2001.
PD
XX 13-AUG-1999; 99EP-00115161.
PF
XX 13-AUG-1999; 99EP-00115161.
PR
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
PI Nueesch J, Rommelaere J;
XX
DR WPI; 2001-212717/22.
DR N-PSDB; AAD02805.
XX
PT Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX
PS Claim 6; Page 30-32; 41pp; English.
XX
CC The present sequence is parvovirus non-structure protein 1 (NS1) variant
CC (T463A). The invention relates to the variants of the parvovirus non-
CC structure protein (NS1) having a shifted equilibrium between the DNA
CC replication and transcription activities, and the cytotoxicity activity.
CC These variants are useful as toxins for treating tumoral diseases. The
CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX
SQ Sequence 672 AA;

Query Match 99.5%; Score 3618; DB 4; Length 672;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 669; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVLGATNWLKEKSNQEVFSFVKNNVQLNGKDIGNWSYKKELODELKSLQ 60
DB 1 MAGNAYSDEVLGATNWLKEKSNQEVFSFVKNNVQLNGKDIGNWSYKKELODELKSLQ 60
QY 61 RGAETTWQSEDMEWETTVDMTKQVIFPDSLVLKVKCLFEVLNTKNIFFGDVNVFVQHEW 120
DB 61 RGAETTWQSEDMEWETTVDMTKQVIFPDSLVLKVKCLFEVLNTKNIFFGDVNVFVQHEW 120
QY 121 GKQDQGHCHVLIGKDFSOAQGWRRQLNYSRWLVTAACNVQLTPAERIKLREIAEDN 180
DB 121 GKQDQGHCHVLIGKDFSOAQGWRRQLNYSRWLVTAACNVQLTPAERIKLREIAEDN 180
QY 181 EWTLLTYKHQTKKDYTKCVLFGNMIAYYFLTKKISTSPRDGGYFLSSDSGKWTNPL 240
DB 181 EWTLLTYKHQTKKDYTKCVLFGNMIAYYFLTKKISTSPRDGGYFLSSDSGKWTNPL 240
```



```
QY 241 KEGERHLVSKLYTDDMRPETVETVTTAQTKEGRIOQTKEVSIKTTLKELVHKRVTSPE 300
Db 241 KEGERHLVSKLYTDDMRPETVETVTTAQTKEGRIOQTKEVSIKTTLKELVHKRVTSPE 300
QY 301 DNMWMPDSYIENMAOPGGGNLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360
Db 301 DNMWMPDSYIENMAOPGGGNLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360
QY 361 TRACRIFAFHGMNVYKCHAIACCVLNRRQGGKRNVLPHFGPASTGKSIIAQAIQAQAVGNVG 420
Db 361 TRTCRIFAFHGMNVYKCHAIACCVLNRRQGGKRNVLPHFGPASTGKSIIAQAIQAQAVGNVG 420
QY 421 CYNAAVNPPFNDCTKNLWVEEAGNFGQVQVNFKAICSGQIRIDQKGGSKQIEPTP 480
Db 421 CYNAAVNPPFNDCTKNLWVEEAGNFGQVQVNFKAICSGQIRIDQKGGSKQIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNEWPMICAWLVK 540
Db 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNEWPMICAWLVK 540
QY 541 NGYQSTMASYCAKWKVPDMSENWABPKVPTPINLLGSARSPFTTPKSTPLSONYALTPL 600
Db 541 NGYQSTMASYCAKWKVPDMSENWABPKVPTPINLLGSARSPFTTPKSTPLSONYALTPL 600
QY 601 ASOLEDLALEPWSPTNTPVAGTAETONTGEGSKACQDQGLSPTWSEIBEDLRACTGAEP 660
Db 601 ASOLEDLALEPWSPTNTPVAGTAETONTGEGSKACQDQGLSPTWSEIBEDLRACTGAEP 660
QY 661 LKXDFSEPLNLD 672
Db 661 LKXDFSEPLNLD 672

RESULT 6
AAP40306
ID AAP40306 standard; protein; 690 AA.
AC
XX AAP40306;
AC
XX 24-OCT-2003 (revised)
DT 12-JAN-1992 (first entry)
XX
XX Sequence encoded by the double-stranded replicative form DNA of porcine
DE parvovirus.
DE
XX Protein envelope; immunogen; vaccine; antigen; epitope.
XX
XX Porcine parvovirus; NADL-2 virulent strain.
XX
XX WO8402847-A.
PN
XX 02-AUG-1984.
XX
XX 19-JAN-1984; 84WO-US0000063.
XX
XX 19-JAN-1984; 83US-00459203.
PR 06-JAN-1984; 84US-00567968.
XX
XX (AMGE-) AMGEN.
PA
XX
XX Fox GM;
PI
XX WPI; 1984-201354/32.
DR N-PSDB; AAN40252.
XX
XX Polypeptide obtd. by recombinant DNA methods - for vaccination against
PT parvovirus infections in man and animals.
XX
XX Claim 6; Table II, Page 33-49; 80pp; English.
PS
XX The inventors claim an immunologically active polypeptide for the
CC development of vaccinal immunity against parvovirus infection. Also
```

```
CC claimed are DNA sequences wholly or partly duplicative of defined
CC sequences. The polypeptides are used in vaccines for conferring
CC protection against parvovirus infections in man and animals. (updated on
CC 24-OCT-2003 to standardise OS field)
XX
SQ Sequence 690 AA;
Query Match 65.2%; Score 2371; DB 1; Length 690;
Best Local Similarity 67.7%; Pred. No. 3.2e-220;
Matches 448; Conservative 70; Mismatches 128; Indels 16; Gaps 6;
QY 2 AGNAYDEVLGATNWLKEKSNQBFVSFVKFNENVLNGDKIGWNSYKKELOEDLSLQOR 61
Db 5 AGNTYSEELKATNWLQDNAQKEAFSYVFTQKVLNKGKEIAWNTYKDTTDAEMINLQOR 64
QY 62 GABTTWDQSEDMETWTTVDKTKQVFI FDSL VKCLFEVLNTKNIFPGDVNWFVQHEWG 121
Db 65 GABTSWDQATDMEWSEIDSLTKRQVLI FDSL VKCLFEGILQKLNLSPSDCYWFQHEHG 124
QY 122 KDQGWCHVLIGGKDFSOAQGWRRQLNVWSRWLVTCNVOLTFAERIKLREIADNE 181
Db 125 QDTGYHCHVLLGGKGLQAMGWRLKQLNMLNSRWLNQCKVPLTPVERIKLRELAEDGE 184
QY 182 WTLTLLYKHKOTKDYTKCVLFGNMIAYYFLTKKISTSPRDGGYFLSDSGWKTNFLK 241
Db 185 WSLTTYTHKQTKKQTKMTHFGNMIAYYFLNKKRKTTE--REHGYLSDSGFWNTFLK 242
QY 242 EGERHLVSKLYTDDMRPETVETVTTAQTKEGRIOQTKEVSIKTTLKELVHKRVTSPE 301
Db 243 EGERHLVSHLFTANKPEITVETVTTAQEVPRGRIQTKKEVSIKTI RDLVNRKCTSIEG 302
QY 302 WMMQPDSEIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPDT 361
Db 303 WSMTPDPSYIENMAQGTGENLLKNTLEITLTARTKTAFDLILEKAPSMPLTFNISNT 362
QY 362 RACRIEAFHGMNVYKCHAIACCVL--NRQGGKNTVLPHGPASTGKSIIAQAIQAQAVGNV 419
Db 363 RTCKIFSMENNVYIKVCHAITCVLKQTRRKKYNSISCHGPASTGKSIIAQHIANLVGNV 422
QY 420 GCYNAANVNPPFNDCTKNLWVEEAGNFGQVQVNFKAICSGQIRIDQKGGSKQIEPT 479
Db 423 GCYNAANVNPPFNDCTKNLWIEEAGNFSNVQVNFKAICSGQIRIDQKGGSKQIEPT 482
QY 480 PVIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNEWPMICAWLV 539
Db 483 PVIMTTNEDITKVRIGCEERPEHTQPIRDRMLNHLTRKLPDGLLEETEWPLICAWLV 542
QY 540 KNGYQSTMASYCAKWKVPDMSENWABPKVPTPINLLGSARSPFTTPKSTPLSONYALTTP 599
Db 543 KKGQATMASYMHWHGNVPDMSKLEBPKMHSPIINTPTDSQIS-TSVKTSPADINYAATP 601
QY 600 LASDLE-DLALPEWSTNTPVAGTAETONTGEGSKACQDQGL---SPTWSEIEEDLRAC 655
Db 602 IQEDLDLALALEPWSPTNTPVAGTAETONTGEGSKACQDQGL---PDSAIRTPSPWSEIETDIRAC 654
QY 656 FG 657
Db 655 FG 656

RESULT 7
AAV58164
ID AAV58164 standard; protein; 550 AA.
XX
XX AAV58164;
AC
XX
XX 07-MAR-2000 (first entry)
DT
DE Adeno associated virus AAV5 Rep68 protein.
XX
XX Adeno associated virus; AAV5; AAV2: inverted terminal repeat; ITR;
KW promoter; Rep protein; capsid protein; regulation; transcription;
KW replication; chromosomal integration; tissue tropism; cellular receptor;
```


CC associated virus 5 (AAV5) particle. AAV5 is a small non-pathogenic virus
CC which relies on a helper virus for replication, in the absence of which
CC the AAV5 genome is integrated into a host chromosome in a locus specific
CC manner. The method provides a way to deliver a nucleic acid to a specific
CC regions, tissues and cell types of the central nervous system comprising
CC inserting the nucleic acid between a pair of AAV inverted terminal
CC repeats or delivering an AAV5 particle containing a vector comprising the
CC nucleic acid. The method is useful for treating brain disorders such as
CC demyelination disease, Alzheimer's disease and Parkinson's disease, and
CC metabolic disorders such as musculoskeletal diseases, cardiovascular
CC disease, cancer and autoimmune disorders, for treating genetic diseases
CC such as cystic fibrosis, alpha-1-antitrypsin, pseudohypoadosteronism,
CC imotile cilia syndrome, and for treating bronchitis, pneumonia,
CC emphysema, and cardiogenic and non-cardiogenic pulmonary oedema. AAV5 is
CC useful for delivering gene that may have a systematic effect like anti-
CC hypertension drugs, insulin, coagulation factors, antibiotics, growth
CC factors and hormones. This is the amino acid sequence of the adeno-
CC associated virus 5 (AAV5) Rep68 protein, one of 4 Rep proteins that
CC regulate replication and transcription of the AAV5 genome, described in
CC the method of the invention
XX
SQ Sequence 550 AA;

Query Match 12.6%; Score 458.5; DB 5; Length 550;
Best Local Similarity 27.4%; Pred. No. 1.2e-34;
Matches 158; Conservative 84; Mismatches 220; Indels 115; Gaps 22;
67 WQSEDMEWETTVDENTKQVIFDLSLVKKCLPEVLNTRKNIFPDGVNMFVQHEWKGDOGW 126
35 WELPPESDLNLTUVE--QPOLTADIRRVFLYE---WNKFSQESKFFVQFEGSEY-F 88
127 HCHVLIGGKDF-SQAQKGKWRRLNVYMSRWLTACNVQLTPAERIKLRBI-AEDNEWVT 184
89 HLHVLVETSGISSMWLGRY-----VSQIRALQV---KVFQGIQPIQNDWA 132
185 LLTYKHQKQKDYTKCVLFGNMIAYFLTKKLISTPPRGGYFLSDSGWKNFLKEGE 244
133 I-----TKVKKGGANKVDSGYIPAYLLPK-----VOPELQAWNTLDE-- 171
245 RHLVSKLYTDDMPETVETVTVAQETKGRIGRIOTKKEVSIKTLKE-----LV 292
172 -YKLAALNLEERKELVAQFLAESQSRQEAASQ--REFSADPVIKSTSKYVALYNLV 228
293 HKRVTSPEDMMQPDYSIEMMAQPGGENLLKNTLICTLTLARTKTAFLDILEKAETSK 352
229 EHGITSKQWIOENQESYLSFNSTGNSRSQIKAALDNATKIMSLTKSAVDY----- 279
353 LTNFSLPD-----TRACIFAFHGN--YV-KVCHACCVLNROGGKNTVLFHGPASTGK 405
280 LVGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGMW---QRSFNKNTVMLYGPATTGK 336
406 SIITAAIAQAVGNVGCYNAANVFPNDCTNKLIVVEAGNFGQOVNQPKAICSGQTR 465
337 TNIAEALHTVPPYGVUNVTNENFPNDVDKMLIWEEGKMTNKVSESALILGSKVR 396
466 IDQKRGSKQIEPTVIMTNENITVVRIGCEERPEHTPIRDRMLNLHLTHPLPGDFGL 525
397 VDQKCKSSVQIDSTPIVTSNTNMVVDGNSSTTFEHOQPLEDRMFKFELTKLPPDFGK 456
526 VDKNEHFMICAWLVKNGYQSTMASYCAKWKGVDPDSE-----NWAEPKVPTPI-- 573
457 ITRQEVKDFPAW-----AKVNOQVPTHEFKVPRELAGTKGAESKLRPLGD 502
574 -----NLGARSAPPT--TPKSTPLSONYA-LTPL 600
503 VMTYSKLEKRLARLFSVPETPRSDVTVDPAPLRPL 539

RESULT 9
AAV58159
ID AAV58159 standard; protein; 610 AA.
XX
AC AAV58159;

XX 07-MAR-2000 (first entry)
XX Adeno associated virus AAV5 Rep78 protein.
XX
XX Adeno associated virus; AAV5; AAV2; inverted terminal repeat; ITR;
XX promoter; Rep protein; capsid protein; regulation; transduction;
XX replication; chromosomal integration; tissue tropism; cellular receptor;
XX gene therapy; neutralising antibody; erythroid progenitor cell;
XX transduction; cancer; genetic disease; Rep78.
XX Adeno-associated virus 5.
XX WO9961601-A2.
XX 02-DEC-1999.
XX 28-MAY-1999; 99WO-US011958.
XX 28-MAY-1998; 98US-0087029P.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Chiorini JA, Kotin RM;
XX WPI; 2000-062707/05.
XX N-PSDB; AA249211.
XX Adeno-associated virus 5 based vectors and particles, useful for gene
XX therapy.
XX Claim 15; Fig 6; 91pp; English.

This sequence represents the Rep78 protein of adeno associated virus type 5 (AAV5). The invention relates to vectors comprising a pair of AAV5 inverted terminal repeats (ITRs) with a promoter between the ITRs. The vector may comprise the viral genome, or subregions thereof, including sequences encoding Rep proteins and capsid proteins, and is encapsidated in an AAV5 particle. The non-structural Rep proteins Rep40 (AAV58613), Rep52 (AAV58168), Rep68 (AAV58164) and Rep78 (AAV58159) are involved in regulation of replication and transcription, in addition to the production of progeny genomes. Rep68 and Rep78 are also associated with the stable integration of the viral genome into human chromosomes. The three types of capsid protein VP1 (AAV58160), VP2 (AAV58161) and VP3 (AAV58162) assemble to form an icosahedral capsid, and differ from each other by the use of alternative splicing and an unusual translation initiation codon (in VP2). AAV5 capsid protein is distinct from AAV2 capsid protein and exhibits different tissue tropism. AAV2 and AAV5 are likely to utilise distinct cellular receptors and are serologically distinct. In a gene therapy application, therefore, AAV5 would allow for transduction of a patient who already possess neutralising antibodies either as a result of natural immunological defence or from prior exposure to AAV2 vectors. The vectors may be useful for transducing erythroid progenitor cells or cells lacking heparin sulphate proteoglycans, which is very inefficient with AAV2-based vectors. The vectors may also be useful for transducing cells with a nucleic acid of interest in order to produce cell lines that could be used to screen for agents that interact with the gene product of the nucleic acid of interest. In addition to transduction of other cell types, transduction of erythroid cells would be useful or the treatment of cancer and genetic diseases which can be corrected by bone marrow transplants using matched donors

SQ Sequence 610 AA;

Query Match 12.6%; Score 458.5; DB 3; Length 610;
Best Local Similarity 27.4%; Pred. No. 1.5e-34;
Matches 158; Conservative 84; Mismatches 220; Indels 115; Gaps 22;

QY 67 WQSEDMEWETTVDENTKQVIFDLSLVKKCLPEVLNTRKNIFPDGVNMFVQHEWKGDOGW 126
Db 35 WELPPESDLNLTUVE--QPOLTADIRRVFLYE---WNKFSQESKFFVQFEGSEY-F 88

127	Qy	127	HCHVILGKDP--SOAQKWRRLQINVYWSRWLTACINWOLTPAERIKLREI--AEDNEWYT 181
89	Db	89	HLHLTVETSGISSMWLGRY-----VSOIRAQLV---KVFGOIEPFIQINDWA 132
185	Qy	185	LLTYKHGQTKDYTKCVLFGNMIAYFYTKKISTSPRDGGYFJSSDSGWTNFKLKEGE 244
133	Db	133	I-----TKVKGGAANKVDSGYIPAYLLPK-----VQPELOWANTNLDE-- 171
245	Qy	245	RHLVSKLYTDDMRPETVTTTAQETKRGRIQTKKEVSIKTKLBE-----LV 292
172	Db	172	-YKLAALNLEERKRLVAQFLAESSORSQFAASQ--REFSADPVIKSTQKYMALWNVLV 228
293	Qy	293	HKRVTSPEDMMWQPDYSYIEMMAQPGGENLLKNLTLEICTLTARTKTAFDLILEKAETSK 352
229	Db	229	EHGITSKQWTOENQESYLSFNSTGNSRSQIIKAALDNATKIMSLYKSAVDY----- 279
353	Qy	353	LTNPSLPD----TRACRIPAFHGN--YV-KVCHAI CCVLNRGGKRNVTLPFHGPASTOK 405
280	Db	280	LVGSSVPDEI1SKNR1WQIFEMNGYDPAYAGS1LYGWC---QRSFNKRNVTWLYGPATTGK 336
406	Qy	406	SIIAQIAQAVGNCVGCYNAANVFPDNCNKLIWBEAGNFGQOVNOFKAICSGOTIR 465
337	Db	337	TNIAEIAHTVPFYGCYVNTNENFPDNCVDKMLIWEEGKMTNKVESAAILGSKVR 396
466	Qy	466	IDQKGKSGQIEPTPVIWMTNENITVWIGCSERPEHTQPIBDRMLN1HLTHLPGDFGL 525
397	Db	397	VDOCKSSVQIDSTPVI1VTSNTNMCCVWDGNS1TTFEHQOPLBDRMFKEPFLTKRLPPDFGK 456
526	Qy	526	VDKNEWPMICAWLVKNGYGSTMASYCAKWKGVDPDWE-----NWAEPKVPPTPI-- 573
457	Db	457	ITKQEVKDFPAW-----AKNQVPVTHFKVPRELAGTKGAEKSLKRPGLD 502
574	Qy	574	-----NLLGSARSPFT--TPKSTPFLSONYA-LTPL 600
503	Db	503	VTNTSYKLEKRLARLSFVPEPRSSDVTVDPA1RLPL 539

RESULT 10

AA197720
ID AAY97720 standard; protein; 610 AA.

XX
AC
AA97720:

DT 19-JUN-2001 (first entry)

Rep protein sequence.

XX Fusion nucleic acid library; Rep protein; tumour cell; apoptosis;
 KW nucleic acid modification enzyme; cell death; decreased cell growth;
 KW protein-protein interaction detection; cell division; cancer therapy;
 KW protein drug discovery; pharmacogenetics.

XX Adeno associated virus 5.

AX
PN
WO200114539-A2.

XX
PD 01-MAR-2001.

18-AUG-2000: 2000WO-US022906.

XX
PR 20-AUG-1999: 99US-0150004P.

PR 02-JUN-2000; 2000US-0209130P.
YY

PA (UYJO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
yy

PI Li M;

DR WPI; 2001-218443/22.

XX
XX

new library of fusion nuclear acids each encoding a Ksp protein, PT
PT recognized by a nucleic acid modification enzyme and a candidate protein,
PT useful for detecting protein-protein interactions, protein drug discovery

or pharmacogenetics.

Disclosure; Fig 21; 106pp; English.

This sequence is the adeno associated virus 5 Rep protein. The invention relates to a library of fusion nucleic acids, each encoding a Rep protein, a candidate protein, a presentation structure, a targeting sequence or a label. The Rep protein is a nucleic acid modification enzyme. The random or directed libraries (including the cDNA libraries) can be introduced into any tumour cell, and peptides identified which by themselves induce apoptosis, cell death, loss of cell division or decreased cell growth. The methods and compositions may also be used to detect protein-protein interactions, protein drug discovery, particularly for protein drugs that interact with targets on cell surfaces, to discover DNA or nucleic acid binding proteins, using nucleic acids as targets, to screen for nucleic acid modification enzymes with decreased toxicity for the host cells, to identify or generate Rep proteins with decreased toxicity, improved enzyme attachment sequences for use in expression vectors and in pharmacogenetic studies. The method is useful in cancer therapy and in killing tumour cells. The methods can be combined with other cancer therapeutics (drugs or radiation) to sensitize cells and thus induce rapid and specific apoptosis, cell death, loss of cell division or decreased cell growth after exposure to a secondary agent

Sequence 610 AA;

ery Match 12.6%; Score 458.5; DB 4; Length 610;

st Local Similarity	27.4%;	Pred. NO.	1.3e-34;		
tches	158; Conservative	84; Mismatches	220; Indels	115; Gaps	22;

67 WDOSEDMEWET"TVDEMTKKOVFI FDSL VKKCLFEVLNTKNI FPGDVNWFVQHEWGKDQGW 126

35 WELPPESDNLTLVE--OPOLTVADRIRRVFLYE---WNKFSKOEKFFVOFEKGSEY-F 88

127 HCHVIIGGKDF-SQAQGKWRRLNLYVSEWLVTCNVOLTPAERIKLREI-AEDNEWVT 184

89 HI:HTI.VETSGISSMVI.GRY-----VSOTRAOI.V---KWFOGIEPOINDWVA 132

185 I.I.TUYNKOTYKKDYTKCVL.EGNMIAVYVEI.TYKKISTSPBPPBDDGGVEI.SSDSCWKTNEI.KEGE 244

[illegible]

245 DIT WGT VERDAMDE MINNEN 3 OORVOND TOEVEZICHT VERREY VE
 246 YU 302

[illegible][illegible][illegible]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible]

280 LVGSSVPEDISNKR1WQIFEMNGIDFAIAGSILIGWC---QKSFNRRKNI VWLIGFAIIION 330

406 SI IAQALAVGNVGCYNAAANVNF PFNDCI NKNLWEEAGNF GQQVNQFKAICS9QIR 463

337 T N I A E A I A H T V P F F Y G C V N W T I N E N F P F N D C V D R M L I W W E E G R M I N K R V E S A K A I L G G S K V R 398

466 IDQKGKSKQIEPTPVIMTTNENITTVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGL 525

397 VDQCKSSVQIDSTPVI VTSNTNMCVVVDGNSTTFEHQQPLEDRMFKFELTKRLPPDFGK 456

526 VDKNEWPMICAWLVKNGYQSTMASYCAKWGKVPDWSE-----NWAEPKVPTPI-- 573

457 ITKQEVKDFFAW-----AKVNQVPVTHEFKVPRELAGTKGAEKSLKRPLGD 502

574 -----NLLGSARSPFT--TPKSTPLSQNYA-LTPL 600

503 VTNTSYKSLEKRALSFVPETPRSSDVTVDPAPLRPL 539

RESULT 11

PT a fusion of nucleic acids encoding NAM enzyme and the compound.

XX Disclosure; Fig 21; 96pp; English.

XX The present invention relates to genetic libraries of nucleic acid/
CC protein (NAP) conjugates comprising a fusion polypeptide (with a nucleic
CC acid modification (NAM) enzyme (S) and candidate compound), an expression
CC vector (with a fusion of nucleic acids encoding the enzyme and candidate
CC protein respectively), an enzyme attachment sequence (EAS; RNA sequence),
CC where the candidate compound and candidate protein are different and EAS
CC and the enzyme are covalently linked. The NAP conjugates are useful in
CC screens to assay binding to target molecules and/or to screen candidate
CC agents for the ability to modulate the activity of the target molecule.
CC They are useful in cancer therapy. Sequences of the invention are also
CC useful to detect protein-protein interaction, in drug discovery, to
CC discover DNA or nucleic acid binding proteins, using nucleic acids as the
CC targets and to screen for NAM enzymes with decreased toxicity for host
CC cells (specifically Rep proteins with reduced toxicity). NAP conjugates
CC are also useful in pharmacogenomic studies, for screening bioactive
CC agents on surface cells, viruses and microbial organisms. They are also
CC useful for screening proteins causing phenotypic changes such as
CC overproduction or inhibition of protein expression, or proteins that
CC alter attachment, infectivity, etc. of the virus. Sequences of the
CC invention are also used in gene therapy. The present sequence is adeno-
CC associated virus (AAV) 5 Rep. (Updated on 29-AUG-2003 to standardise OS
CC field)

XX SQ Sequence 610 AA;

Query Match 12.6%; Score 458.5; DB 5; Length 610;

Best Local Similarity 27.4%; Pred. No. 1.5e-34;

Matches 158; Conservative 84; Mismatches 220; Indels 115; Gaps 22;

QY 67 WQOSEDMEWTTVDKTKQVFFDLSLVKKCLFEVLNTKNIIPGCVNMFVQHEWGXQDGW 126
DB 35 WELPPSDMLNLTLVE--QPOLTVDADRIRRVFLVE---WNKFSKQESKFFVQFEGSEY-F 88
QY 127 HCHVLIGGKDF-SQAQGWRRQLNVYWSRWLTACNVQLTPAERIKLREI-AEDNEWVT 184
DB 89 HLHLTVETSGISSWVLGRY-----VSQIRALV---KVVFQIEQINDWA 132
QY 185 LLTYKHQTKDYTKVLFNGMIAYFLTKKKISTSPRGGYFLSDSGWKNFLKEGE 244
DB 133 I-----TKVKKGGANKVVDGYPAYLLPK-----VOPELQAWTNLDE-- 171
QY 245 RHLVSKLYTDDMPETVETVTTAQETKRGRIOTKKEVSIKTLKE-----LV 292
DB 172 -KLAALNLEERKRLVAQFLAESQSQEASQ--REFSADPVIKSTSKYVALVNLV 228
QY 293 HKEVTSPEWMMQPSYIEMMAQPGGENLLKNTLEICTLTARTTAPDLILEKATSK 352
DB 229 EHGITSEKQMIQENQESYLSFNSTGNSRSQIKAALDNATKIMSITSKAVDY----- 279
QY 353 LTNFSLPD-----TRACHIPAFHGN--YV-KVCHAIICCVLNROGGKRNVTFLFGPASTGK 405
DB 280 LVGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGWC---QRSFNKENTVWLYGPATTKG 336
QY 406 SIIAQATAQAVNGVGVNAANVFPNDCKNKLIVVEBAGNFQQVQKALCSGOTIR 465
DB 337 TNIAEAIAHTVPFGCVNVTNEPFPNDCVCKMLIMWEEBGMKTNKVESAKAILGSKVR 396
QY 466 IDQKGKSGKQIEPTPVIMTNNITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGL 525
DB 397 VDQKSSSVQIDSTPVIIVTNTNMCVVVDGNTSTTFEHQPLEDRMFKFELTKLPDPFGK 456
QY 526 VDKNEWPMICAWLVNGYQSTMASYCAKWKGVDPWSE-----NWAEPKVPTPI-- 573
DB 457 ITRQEVKDFFAW-----AKVQVPTVTFEYKVPRELAGTKGAESKLKRLPGD 502
QY 574 -----NLGARSRPPT--TPKSTPLSONYA-LTPL 600
DB 503 VWTYSKLEKRLARLSFVPTPRSSDVTVDPAFLRPL 539

RESULT 13

AAE28644

ID AAE28644 standard; protein; 610 AA.

XX AAE28644;

AC AAE28644;

XX 29-AUG-2003 (revised)

DT 27-DEC-2002 (first entry)

XX

DE Adeno-associated virus 5 Rep protein.

XX

KW Nucleic acid modification enzyme; NAM; enzyme attachment sequence; EAS;
KW protein design automation; PDA; cancer; protein-protein interaction;
KW infection; gene therapy; Rep protein.

XX

OS Adeno associated virus; 5.

XX

FN WO200268453-A2.

XX

PD 06-SEP-2002.

XX

PF 19-FEB-2002; 2002WO-US004853.

XX

PR 22-FEB-2001; 2001US-00792629.

XX

PA (XENC-) XENCOR INC.

XX

PI Li M, Dahiyat BI;

XX

DR WPI; 2002-691653/74.

DR

DR N-PSDB; AAD46138.

XX

PT Generating a library of fusion nucleic acids for treating cancer or
PT infection, or detecting protein-protein interaction, comprises providing
PT computationally-derived library of candidate protein sequences and
PT expression vectors.

XX

PS Disclosure; Page 180-182; 246pp; English.

XX

CC The present invention relates to a novel method of generating a library
CC of fusion nucleic acids. The method involves providing a computationally-
CC derived library of candidate protein sequences and creating a library of
CC expression vectors containing a fusion nucleic acid having a sequence
CC encoding a nucleic acid modification (NAM) enzyme and a sequence encoding
CC a candidate protein sequence from the library and an enzyme attachment
CC sequence (EAS) that is recognised by the NAM enzyme. The invention also
CC relates to the use of a variety of computation methods including protein
CC design automation (PDA). The method is useful in generating and screening
CC fusion nucleic acids that may be used in treating cancer or infections,
CC in detecting protein-protein interactions, discovery of DNA or nucleic
CC acid binding proteins, protein drug discovery, screening for NAM enzymes
CC with decreased toxicity to the host cells and NAM enzyme/EAS pairs with
CC increased affinity or in pharmacogenetic studies. The invention is also
CC used in gene therapy. The present sequence is Adeno-associated virus 5
CC Rep protein. This sequence is used to illustrate the method of the
CC invention. (Updated on 29-AUG-2003 to standardise OS field)

XX SQ Sequence 610 AA;

Query Match 12.6%; Score 458.5; DB 5; Length 610;

Best Local Similarity 27.4%; Pred. No. 1.5e-34;

Matches 158; Conservative 84; Mismatches 220; Indels 115; Gaps 22;

QY 67 WQOSEDMEWTTVDKTKQVFFDLSLVKKCLFEVLNTKNIIPGCVNMFVQHEWGXQDGW 126

DB 35 WELPPSDMLNLTLVE--QPOLTVDADRIRRVFLVE---WNKFSKQESKFFVQFEGSEY-F 88

QY 127 HCHVLIGGKDF-SQAQGWRRQLNVYWSRWLTACNVQLTPAERIKLREI-AEDNEWVT 184

DB 89 HLHLTVETSGISSWVLGRY-----VSQIRALV---KVVFQIEQINDWA 132

QY 185 LLTYKHQTKDYTKVLFNGMIAYFLTKKKISTSPRGGYFLSDSGWKNFLKEGE 244


```
Db 133 i-----TKVKKGGANKVDSGYIPAYLLPK-----VQPELQAWNTLDE-- 171
Qy 245 RHLVSKLYTDDMRPETVETTTTAQETKRGRIOTKKEVSIKTLKE-----LV 292
Db 172 -YKLAALNLEERKRLVAQFLAESSQRSQAASQ--REFSADPVIKSTOKYMAVNLV 228
Qy 293 HKRVTSPEDDMMQPDYSIEMMAQPGENLLKNTLICTLTARTKTAFDLILEKAETSK 352
Db 229 EHGITSKQWIOENQESYLSFNSTGNSRSQIKALDNATKIMSLTKSAVDY----- 279
Qy 353 LTNFSLPD----TRACRIFAFHGMN--YV-KVCHAICCVLNROGGKNTVLFHGPASTGK 405
Db 280 LVGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGC--QRSFNKNTVWLYGPATTGK 336
Qy 406 SIITAQIAQAVNGVGCYNAANVPFNDCNTKNLIWVEAGNFGQOVNOFKATCSQTIR 465
Db 337 TNIAEAIAHTVPYGCYVNTNENFPFNDKMLIWWEEGKMTNKVVESAKALGSKVR 396
Qy 466 IDQKGKSKQIEPTPVIMTNTNITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGL 525
Db 397 VDQCKSSVQIDSTPVIIVTSNTNMCVVVDGNSSTTFEHQOPLDRMFKFELTKRLPPDFGK 456
Qy 526 VDKNEWPMICAWLVKNGYQSTMASYCAKWKGVDPDSE-----NWAEPKVPTPI-- 573
Db 457 ITKQEVKDFEAW-----AKNQVPVTHEFKVPRELAGTKGAESLKRPLGD 502
Qy 574 -----NLLGSARSPT--TPKSTPLSONYA-LTPL 600
Db 503 VTNYSKLEKRLARLSFVPETPRSSDVTVDPAERPL 539

RESULT 14
AAE26940
ID AAE26940 standard; protein; 610 AA.
AC AAE26940;
XX
DT 13-DEC-2002 (first entry)
XX
DE Adeno associated virus 5 Rep protein.
XX
KW Prokaryotic library; candidate protein; nucleic acid modification; NAM;
KW enzyme attachment sequence; EAS; clinical pharmacology; chemical sensor;
KW enzymology; cosmetic research; toxic; environmental safety assessment;
KW nutrient biology; Rep protein.
XX
OS Adeno associated virus.
XX
PN WO200266653-A2.
XX
PD 29-AUG-2002.
XX
PF 14-DEC-2001; 2001WO-US049058.
XX
PR 14-DEC-2000; 2000US-0256163P.
XX
PA (XENC-) XENCOR INC.
XX
PI Li M, Liu Y;
XX
DR WPI; 2002-667068/71.
DR N-PSDB; AAD44600.
XX
PT New library of prokaryotic pET-24a expression vectors, host cells or
PT nucleic acid/protein conjugates, useful for screening candidate proteins
PT and their nucleic acids or modification enzymes for pharmacogenetic
PT analysis.
XX
PS Disclosure; Fig 21; 127pp; English.
XX
CC The invention relates to methods and compositions for the construction of
CC prokaryotic libraries expressing candidate proteins and the use of these
```

```
CC libraries to identify candidate proteins and the nucleic acids encoding
CC them. The invention provides a library of prokaryotic pET-24a vectors
CC comprising a fusion nucleic acid consisting of a nucleic acid encoding a
CC nucleic acid modification (NAM) enzyme or a candidate protein, or a
CC nucleic acid having a T7 promoter operably linked to the NAM enzyme or
CC the candidate protein, and an enzyme attachment sequence (EAS) recognised
CC by the NAM enzyme. The library is used for identifying candidate proteins
CC and nucleic acids encoding these proteins, in screening for NAM enzymes
CC with decreased toxicity for the host cells, or in identifying novel or
CC improved EASs, which may be used for understanding cellular processes or
CC any subsequent therapeutic or toxic activities. The nucleic acid/protein
CC (NAP) conjugates are useful in diagnostic assays and in research
CC including clinical pharmacology, functional genomics, pharmacogenomics,
CC agricultural chemicals, environmental safety assessment, chemical sensor,
CC nutrient biology, cosmetic research or enzymology. These may also be used
CC in vitro screening techniques and in assays with target molecules. The
CC present sequence is Adeno associated virus 5 Rep protein used in the
CC invention
XX
SQ Sequence 610 AA;
```

```
Query Match 12.6%; Score 458.5; DB 5; Length 610;
Best Local Similarity 27.4%; Pred. No. 1.5e-34;
Matches 158; Conservative 84; Mismatches 220; Indels 115; Gaps 22;
```

```
Qy 67 WQOSEDMEVETTVDEMTKQVFIQFDSLKKLPEVLNTKNIFPGDVNWFVQHEWKGDKQGW 126
Db 35 WELPPESDLNLTVE--QPQLTVADRIRRVFLVE---WNKFSQESKSFVQFKGSEY-F 88
```

```
Qy 127 HCHVLIGGKDP-SQAQGWRRQLNLYVWSRWLTACNVOLTPAERIKLRI-AEDNEWVT 184
Db 89 HLHVLTVETSGISSMWLGY-----VSQIRAOQLV---KVFQGIETPOINDMVA 132
```

```
Qy 185 LLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPRDGGYFLSDSGKWNFLKEGE 244
Db 133 i-----TKVKKGGANKVDSGYIPAYLLPK-----VQPELQAWNTLDE-- 171
```

```
Qy 245 RHLVSKLYTDDMRPETVETTTTAQETKRGRIOTKKEVSIKTLKE-----LV 292
Db 172 -YKLAALNLEERKRLVAQFLAESSQRSQAASQ--REFSADPVIKSTOKYMAVNLV 228
```

```
Qy 293 HKRVTSPEDDMMQPDYSIEMMAQPGENLLKNTLICTLTARTKTAFDLILEKAETSK 352
Db 229 EHGITSKQWIOENQESYLSFNSTGNSRSQIKALDNATKIMSLTKSAVDY----- 279
```

```
Qy 353 LTNFSLPD----TRACRIFAFHGMN--YV-KVCHAICCVLNROGGKNTVLFHGPASTGK 405
Db 280 LVGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGC--QRSFNKNTVWLYGPATTGK 336
```

```
Qy 406 SIITAQIAQAVNGVGCYNAANVPFNDCNTKNLIWVEAGNFGQOVNOFKATCSQTIR 465
Db 337 TNIAEAIAHTVPYGCYVNTNENFPFNDKMLIWWEEGKMTNKVVESAKALGSKVR 396
```

```
Qy 466 IDQKGKSKQIEPTPVIMTNTNITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGL 525
Db 397 VDQCKSSVQIDSTPVIIVTSNTNMCVVVDGNSSTTFEHQOPLDRMFKFELTKRLPPDFGK 456
```

```
Qy 526 VDKNEWPMICAWLVKNGYQSTMASYCAKWKGVDPDSE-----NWAEPKVPTPI-- 573
Db 457 ITKQEVKDFEAW-----AKNQVPVTHEFKVPRELAGTKGAESLKRPLGD 502
```

```
Qy 574 -----NLLGSARSPT--TPKSTPLSONYA-LTPL 600
Db 503 VTNYSKLEKRLARLSFVPETPRSSDVTVDPAERPL 539
```

```
RESULT 15
ABU64865
ID ABU64865 standard; protein; 610 AA.
XX
AC ABU64865;
XX
DT 14-MAY-2003 (first entry)
```

XX	Rep protein sequence from adeno-associated virus 5.
DE	
XX	
XX	Rep protein; capture probe; expression vector;
KW	nucleic acid protein conjugate; NAP; enzyme attachment sequence; EAS;
KW	biochip; gene expression profiling; mutation detection; Rep68; Rep78;
XX	nonstructural protein; NS1; major coat protein; U94.
XX	
XX	Adeno-associated virus 5.
OS	
XX	
PN	US2002172968-A1.
XX	
PD	21-NOV-2002.
XX	
PF	19-FEB-2002; 2002US-00080376.
XX	
XX	22-FEB-2001; 2001US-00792630.
PR	
XX	(LIUH/) LIU H.
PA	(DAHI/) DAHIYAT B I.
PA	(LIHM/) LI M.
PA	
PI	Liu H, Dahiayat BI, Li M;
XX	
XX	WPI; 2003-310986/30.
DR	N-PSDB; ABX96669.
XX	
XX	New composition comprising a substrate consisting of an array of capture
PT	probes hybridized to an expression vector or to a nucleic acid protein
PT	conjugate, useful for diagnostic test, gene expression profiling or
PT	mutation detection.
XX	
PS	Disclosure; Fig 21; 125pp; English.
XX	
CC	The invention relates to a composition comprising a substrate comprising
CC	an array of capture probes hybridized to an expression vector or to a
CC	nucleic acid protein conjugate. The capture probes are hybridized to an
CC	expression vector or to a nucleic acid protein (NAP) conjugate. The
CC	vector comprises: (a) a fusion nucleic acid; (b) a capture sequence; and
CC	(c) an enzyme attachment sequence (EAS). The NAP conjugate comprises: (a)
CC	a fusion polypeptide; and (b) an expression vector. The fusion nucleic
CC	acid comprises a nucleic acid encoding the NAP enzyme or candidate
CC	protein. The fusion polypeptide comprises a Rep and candidate protein.
CC	The EAS and NAP enzyme are covalently attached. Also included are
CC	detecting the presence of a target analyte in a sample, making biochips,
CC	and making NAP conjugates. The composition is useful for diagnostic
CC	applications, gene expression profiling or mutation detection. The
CC	present sequence represents a viral Rep (or related protein e.g. Rep68,
CC	Rep78, nonstructural protein, NS1, major coat protein or U94 protein) for
CC	use in the composition of the invention
XX	
XX	Sequence 610 AA;
SQ	

[illegible]

Search completed: January 22, 2005, 03:35:22
Job time : 103 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:27:50 ; Search time 25 Seconds
(without alignments)
1782.628 Million cell updates/sec

Title: US-10-069-056-10
Perfect score: 2637
Sequence: 1 MAGNAYSDEVLGATNWLKEK.....RACFGAEPLKKDFSEPLNLD 672

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	437.5	12.0	536	4	US-09-532-594B-10
2	437.5	12.0	623	4	US-09-532-594B-2
3	437.5	12.0	623	4	US-09-532-594B-11
4	431.5	11.9	546	4	US-09-807-802A-7
5	431.5	11.9	623	4	US-09-807-802A-2
6	431.5	11.9	623	4	US-09-807-802A-5
7	430.5	11.8	312	4	US-09-532-594B-8
8	430.5	11.8	399	4	US-09-532-594B-9
9	423.5	11.6	322	4	US-09-807-802A-11
10	423.5	11.6	399	4	US-09-807-802A-9
11	144	4.0	276	3	US-09-171-461-3
12	144	4.0	276	4	US-09-970-711-3
13	121.5	3.3	647	3	US-09-300-909-17
14	112	3.1	589	4	US-09-107-532A-6086
15	111	3.1	564	3	US-09-134-001C-2897
16	105	2.9	481	4	US-09-252-991A-30905
17	102	2.8	1170	4	US-09-638-524A-2
18	101	2.8	715	4	US-09-248-796A-20602
19	101	2.8	1120	4	US-09-792-024-95
20	100.5	2.8	2409	6	5180808-2
21	100	2.7	605	4	US-09-107-532A-6077
22	100	2.7	1503	4	US-09-677-046A-6
23	100	2.7	1509	4	US-09-677-046A-2
24	99	2.7	919	4	US-09-248-796A-19408
25	98.5	2.7	657	3	US-09-300-909-16
26	98	2.7	1343	4	US-09-270-767-31792
27	98	2.7	1343	4	US-09-270-767-47009

28	97.5	2.7	556	4	US-09-248-796A-18763	Sequence 18763, A
29	97.5	2.7	3248	1	PCT-US95-700-1	Sequence 1, Appl1
30	97.5	2.7	3248	5	PCT-US95-16216-1	Sequence 1, Appl1
31	96.5	2.7	957	4	US-09-489-039A-11233	Sequence 11233, A
32	96	2.6	291	4	US-09-107-532A-6390	Sequence 6390, Ap
33	96	2.6	422	3	US-09-134-001C-3034	Sequence 3034, Ap
34	95.5	2.6	441	4	US-09-540-236-2983	Sequence 2983, Ap
35	95	2.6	590	4	US-09-134-000C-4755	Sequence 4755, Ap
36	95	2.6	1307	1	US-08-395-246C-2	Sequence 2, Appl1
37	94	2.6	370	4	US-09-583-110-3405	Sequence 3405, Ap
38	94	2.6	1477	4	US-09-206-942-71	Sequence 71, Appl
39	94	2.6	1767	4	US-09-583-110-3975	Sequence 3975, Ap
40	94	2.6	2138	4	US-09-583-110-5274	Sequence 5274, Ap
41	94	2.6	15281	2	US-08-471-119A-2	Sequence 2, Appl1
42	93.5	2.6	506	4	US-09-540-236-2242	Sequence 2242, Ap
43	93.5	2.6	692	4	US-09-248-796A-19136	Sequence 19136, A
44	93.5	2.6	821	3	US-09-422-869-24	Sequence 24, Appl
45	93.5	2.6	821	4	US-09-538-092-972	Sequence 972, App

ALIGNMENTS

RESULT 1
US-09-532-594B-10
; Sequence 10, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safar, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; OTHER INFORMATION: synthetic construct
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 Rep protein 68
US-09-532-594B-10

Query Match	12.0%	Score 437.5;	DB 4;	Length 536;
Best Local Similarity	27.7%	Pred. No. 1.7e-36;		
Matches	157;	Conservative	75;	Mismatches 232; Indels 103; Gaps 19;
QY	63	AETWQSDMEWETTVDEMTKQVIFDLSLVKKCLFEVLTKNIFPGDVNFWVQHEWGK	122	
DB	31	AEKWEELPPDSMDLNLIE--QAPLTVAEKQLQREFLVEMRRVSK--APEALFFVQPEKG-	85	
QY	123	DOGMWCHVLIG-----GKDFSOAQGQKWRQLNVYMSRWLTACNVQLTPAERIKL	173	
DB	86	DSYFHLHLVETGVKSMVGVRSIQEK-----LVTRIYRGVEP-----	126	
QY	174	RETAEDNEWTLTYKHQ--TKQDYTKCVLFGNMIAIYELTKKKIKISTSPRGGGYFLSS	231	
DB	127	-----QLPNPNFAVTKTRNGAGGNKVDDC-----YIPNYLLPK-----TQP-----	163	
QY	232	DSGMK--TNF-----LKEGERHLVSKLYTDDMRPEVETVTTTAQETKRG-	275	
DB	164	ELQWATNMDDQYISACINLAERKRLVAQHL-----THVSOTQEQGNQNPNSD	212	
QY	276	---IQTKKEVSIKTKLKLVLHKRVTSPEWMMQPDYSIEMMAQPGGNNLLKLTLEICTL	332	
DB	213	APVIRKTSARYNELVGLVDRCITSEKQWIQEDQASYSFNAASRSRQIKKALDNASK	272	

393	NTVLFHGPGASTGKSI	IIAQIAQAQVGNVGCYNAAVNFPNDCTNKMILIWBEAGNPGQOV	452
Qy			
328	NTIWLFGPATTTGNTIAE	IAHAAPFYGCYNVTNENFPNDCCVDMKVIWEEEGKMTAKVV	387
Db			
453	NQKKAICSQOTIRIDOKG	SKSIOETPTVMTTNNENITVVRIGCERPEHTOPIRDLNL	512
Qy			
388	ESAKALGGSKVRVDQK	SSAQIDPTPVLIVTSNTNMCAVTDGNSTTTFHQDPLQDRMPK	447
Db			
513	ILHTLHPGDFGLVDKNE	WPMICAWLWQNGYQSTMAASYCAKMG--KVPDWSS--NWAEPKV	569
Qy			
448	FELTKRLEHDFGKVTQ	KEVQDKPFRWADSHVTEVTFEYVRKGGARKPAPNDADISEPKX	507
Db			
570	PTPINLIGSARSPTTPK	STPLSONYA	596
Qy			
508	ACP-----SVAOPST	SDAEPV--DVA	527
Db			

RESULT 4
US-09-807-802A-7
; Sequence 7, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN-031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 546
; TYPE: PRT
; ORGANISM: AAV-1
; US-09-807-802A-7

Query Match	11.94;	Score	431.5;	DB	4;	Length	546;
Best Local Similarity	27.54;	Pred.	No. 7.4e-36;				
Matches	156;	Conservative	64;	Mismatches	224;	Indels	123;
Gaps	19;						
Qy	63	AETTWDQSDMEWE-----	TTVDMTKKQVFIFDLSLVKKCLPEVLNTKNIFFGDVNW	114			
Db	31	AKEWELPPDSMDLNLIEQAPITVAEKLODFLVQWRRVSKA-----	PEALP	78			
Qy	115	FVQHEWGKDGWCHVLIIG-----	GKDFSQAQGWKR-----	ROLNVYWRSLWLTACN	162		
Db	79	FVQPEKGESY-FHLHILVETGVKSMVLGRFLSQIRDKLVTQTYRGIEPTLPNWEFAVT--	135				
Qy	163	VQLTPAERIKLREIAEDNEWVLLTYKHQKTKDYTKCVLFGNMIAYVLTKKKISTSP	222				
Db	136	-----KTRNGAGGNGKVDECY-----	IPNYLPK-----	TQB-	163		
Qy	223	RDGGYFLSSDSGWK-TNF-----	LKEGERHLVSKLYTDMDRPETVETTVAQETK	272			
Db	164	-----ELQAWTWNMEEYISACLNLAERKLVAQHL-----	THVSQTQEQN	203			
Qy	273	RGRIQTKEYSI---KTTLKE-----	LVHKRTVSPEDWMMQPDYSYIEMMAQPGGENLL	323			
Db	204	KENLNPNSDAPVIRSKTSABRYMELVGLVDRGITSEKQWIQEDQASYSIFSNAAASRSQI	263				
Qy	324	KNTLIEICTLTLARTKTAFDLILEKAETSKL-TNFSLPDTRACRIFAFHGWNVYKVCCHAIC	382				
Db	264	KAALDNAGKTMALTKSAPDYLVGAPPADIKTN-----	RIYKILENGVEPAYAGSVFL	317			
Qy	383	CVLNRQGGKRNVLV.FHGPASTGKSIYIAQATAQAVGNVGCVNAANVNFPPNDCNKLW	442				
Db	318	GWAOKRFGKENTLWFGPATTTGKNTAEATAHVAVPFGVGNWNTNFPPNDCVCKKVI	377				

```

Qy 443 EENAGFEQQVQVQKAIKCSGTTRIDQKGGKSKQIEPTPIMTNNENITVVRIGCERPEH 502
Db 378 EEGMTAKVVESSAKAIIIGSKVRVDQKCKSSAQIDPTPIVITSNTMCAVIDGNSTTFEH 437
Qy 503 TQPIRDRMLNIHLTHLPGDVGGLVDKNWPMICAWLVKXGYSOTMAASYCAKWG----KVP 558
Db 438 QQFLQDRMFKFELTRLEHDFGKVTQEVKEPFRWAQDHVTEVAHEFPYVRKGANKRPAP 497
Qy 559 DWSENMAEPKVPTPINLGSARSPFTT 585
Db 498 DDADK-SEPKRACP-----SVADPSTS 518

RESULT 5
US-09-807-802A-2
; Sequence 2, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNVPN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 9233
; TYPE: PRT
; ORGANISM: AAV-1
; US-09-807-802A-2

```

Query Match	11.9%	Score	431.5	DB	4	Length	623
Best Local Similarity	27.5%	Pred. No.	9.4e-36				
Matches	156	Conservative	64	Mismatches	224	Indels	123
Gaps	19						
Qy	63	AEITWDSQSDMEWE-----TTVDENMKQVFI	FDLSLVKKCLFVLNKNIFPGDVNW	114			
Db	31	AEXEWELPPSDMDLNLIEOAPLTVAEKIQ	DFLVQWRRVSKA-----PEALF	78			
Qy	115	FVQHEMGKQGWCHVHLIG-----GKDFSQAQCKWNR---	ROLNVVSRWLVTAGN	162			
Db	79	FVQPEKGESY-FHLHILVETTVGKSMVLGR	FLSQIRDKLVTQIYRGIEPLPNWPAVT--	135			
Qy	163	VQUTPAERIKLRIBAEEDNEWVTLTYTKHKQTK	KDYTKCVLFGNMIAYYFLTKKIKSTSP	222			
Db	136	-----KTRNGAGGNNKVDECY-----	IPNLLPK-----TQP-	163			
Qy	223	RDGGYFLUSSQGWK-TNF-----LKEGRHLVSKLY	TTDDMRPETVETTVAQETK	272			
Db	164	-----BLQAWNTMEEYISACLNLAERKELVAQHL---	THVSQTQEQN	203			
Qy	273	RGRIQTKKYSI---KTTLKE-----LVHKRYTSP	EDMMQPDSEYIEMWAPGGENLL	323			
Db	204	KENLNPNSDAPVTRTSARYMELVGNLVDRGITS	SEKQWIEDQOQSYISFNAAASRSQI	263			
Qy	324	KNTLIECTLTLARTKAFDLILEKAETSKL-TNFSLP	DFTRACHIFAFHGNNYKVCVCHAIC	382			
Db	264	KAALDNAGKIMALTKSAPDVLVGPAPADIKTN-----	RIYRILELNGEPAYAGSVFL	317			
Qy	383	CVLNROGGKNTVLTHGPASTGKSIIAQAIQAQ	AVGNVCVNAANVPFPDNCNTNKLIVW	442			
Db	318	GWAQKRGKRNITWLFGPATGTGNTIAEAIAHA	VYPYGCVNTNENFPFNDCCVDKMIW	377			
Qy	443	BEAGNFQOQVNPQKATCSGGTIRIDOKGSKQI	EPTVPIMTNENITVVRIGCEERPEH	502			
Db	378	BEGKNTAKVYESAKAILGGSKVRDQCKSSAQID	PTPVLVTSNTNCAVIDGNSTTTFH	437			

Qy	503	TOPIRDLMLNIHLTHHLPGDECLVDKNEPNCI	AWLVKNGVQSTMASYCAKWG----	KVP	558
Db	438	QQQLQDMMKFELTRRLHDFGVKTVQEKVEFFR	MAQDHVTEVAHEFYVKKGGANKRPAP	497	
Qy	559	DWSNWAEPKVPTPINLIGSARSPETT	585		
Db	498	DDADK--SEPKRACP-----SVADPETS	518		

```

RESULT 6
US-09-807-802A-5
; Sequence 5, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Willson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 623
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-5

```

QY 559 DWSNWAEPKVPPTPINLGSARSPFTT 585
| : : | | | | | | | | :
DB 498 DDADK-SEPKRACP-----SVADPSTS 518

```

RESULT 7
US-09-532-594B-8
; Sequence 8, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotlin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AA4V VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 312
; TYPE: PrT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence
; NAME/KEY: misc feature
; OTHER INFORMATION: AA4V Rep protein 40
US-09-532-594B-8

```

RESULT 8
US-09-532-594B-9
; Sequence 9, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AA4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.025202
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22


```

; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6469524e =
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 Rep protein 52
US-09-532-594B-9

Query Match      11.8%; Score 430.5; DB 4; Length 399;
Best Local Similarity 35.6%; Pred. No. 5.3e-36;
Matches 110; Conservative 45; Mismatches 139; Indels 15; Gaps 5;

QY 291 LVHKRVTSPEDEMMQPDSDYIEMMAQPGGKNTLKTLEICTLTARTKTAFDILEKAET 350
DB 7 LVDRGITSEKQIQEDQASYISFNAASRSQIKAAALDNAGKIMALTAKSAPDYLVGPP 66
QY 351 SKL-TNFSLPDTRACRIFAFHGMNVYKVAICCVLNROGGKNTVLFHGPASTGKSI 409
DB 67 ADIKTN-----RIYRILELNGYEPAYAGSVFLGWAQKRFGRKNTIWLFGPATTK 120
QY 410 QATAQAVGNVGCYNAANVPFNDCTNKNLIWVEEAGNFGQVQVQPKAICSGQTIR 469
DB 121 EALAHAVPFYGCNVNTNENFPFNDCTNKNLIWVEEAGNFGQVQVQPKAICSGQ 180
QY 470 GKSGKOIEPTPVIMTNNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGD 529
DB 181 CKSSAQIDPTPIVTSNTNCAVIDGNSITTFEHOQPLQDRMFKFELTRLEHDFG 240
QY 530 EWPMICAWLVKNGYQSTMASYCAKMG-----KVPDWSNNAEPKVPPTPINL 595
DB 241 EVKEFFRWAQDHVTEVAHEFYVRKGGANKRPAPDDADK-SEPKRACP----- 294

RESULT 10
US-09-807-802A-9
; Sequence 9, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilison, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 399
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-9

Query Match      11.6%; Score 423.5; DB 4; Length 399;
Best Local Similarity 36.3%; Pred. No. 2.9e-35;
Matches 109; Conservative 38; Mismatches 136; Indels 17; Gaps 5;

QY 291 LVHKRVTSPEDEMMQPDSDYIEMMAQPGGKNTLKTLEICTLTARTKTAFDILEKAET 350
DB 7 LVDRGITSEKQIQEDQASYISFNAASRSQIKAAALDNAGKIMALTAKSAPDYLVGPP 66
QY 351 SKL-TNFSLPDTRACRIFAFHGMNVYKVAICCVLNROGGKNTVLFHGPASTGKSI 409
DB 67 ADIKTN-----RIYRILELNGYEPAYAGSVFLGWAQKRFGRKNTIWLFGPATTK 120
QY 410 QATAQAVGNVGCYNAANVPFNDCTNKNLIWVEEAGNFGQVQVQPKAICSGQTIR 469
DB 121 EALAHAVPFYGCNVNTNENFPFNDCTNKNLIWVEEAGNFGQVQVQPKAICSGQ 180
QY 470 GKSGKOIEPTPVIMTNNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGD 529
DB 181 CKSSAQIDPTPIVTSNTNCAVIDGNSITTFEHOQPLQDRMFKFELTRLEHDFG 240
QY 530 EWPMICAWLVKNGYQSTMASYCAKMG-----KVPDWSNNAEPKVPPTPINL 595
DB 241 EVKEFFRWAQDHVTEVAHEFYVRKGGANKRPAPDDADK-SEPKRACP----- 294

US-09-807-802A-11
; Sequence 11, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilison, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 322
; TYPE: PRT
; ORGANISM: AAV-1
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (943)..(944)
; OTHER INFORMATION: minor splice site
US-09-807-802A-11

Query Match      11.6%; Score 423.5; DB 4; Length 322;
Best Local Similarity 36.3%; Pred. No. 2e-35;
Matches 109; Conservative 38; Mismatches 136; Indels 17; Gaps 5;

```

```
; Sequence 3, Application US/09171461
; Patent No. 6335016
; GENERAL INFORMATION:
; APPLICANT: Baker, Adam
; APPLICANT: Cotten, Matthew
; APPLICANT: Chiocci, Susanna
; APPLICANT: Kurzbauer, Robert
; APPLICANT: Schaffner, Gotthold
; TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus
; FILE REFERENCE: 0652.1800000
; CURRENT APPLICATION NUMBER: US/09/171,461
; CURRENT FILING DATE: 1999-01-12
; EARLIER APPLICATION NUMBER: PCT/EP97/01944
; EARLIER FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 276
; TYPE: PRT
; ORGANISM: CELO Virus
; FEATURE:
; OTHER INFORMATION: Position: 1999..2829 /note=ORF2
US-09-171-461-3

Query Match      4.0%; Score 144; DB 3; Length 276;
Best Local Similarity 24.1%; Pred. No. 3.7e-06;
Matches 64; Conservative 41; Mismatches 133; Indels 28; Gaps 9;

QY 286 TTLKELVHKRVTSPEWMMQPDY--IEMMAQPGGENLLKNTLEICTTLTARTKTAFDL 343
DB 10 TLVHALIDRGVSRQWQVDPAYQFYHRSKORGFK--VRHILRDVIRHMCWSRTLLDY 67
QY 344 ILEKAETSKLTNFSLPD---TRACRIFAFHGMNYYKVKCHAI CCVLNRQGGKENTVLFHG 399
DB 68 MSSASTPS-----PDDVLNPLYLQLLCNGYNPAVVGTTALIRWAGHQ-SNRTVWIRG 119
QY 400 PASTGKSIITAAIAQAVNGVGCYNAANVPFNDCTNKNLIWVEEAGNFQGVNQPKAIC 459
DB 120 TPMSGAPYLAQAIAYCSPVGSVDNRKSNPFEGCPDSLFWMDGGVYVDCCVGLVKQVF 179
QY 460 SGOTIRIDQK-----KGSQKIEPTVMTNENITVVRIGCEE-RPEHTOPIRDRMLNIH 514
DB 180 RGEHVLPPEGLRGNPCSELFRTPVLMYSQADICMTRLRSGLSAEHAVALGRDCMYLIR 239
QY 515 LTHLPGDF---GLVDKNEWPMICAW 537
DB 240 LTE----DFDCAGGISCADVKQFVAW 261

RESULT 13
US-09-300-909-17
; Sequence 17, Application US/09300909
; Patent No. 6306580
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PREPARATION OF HUMAN PAPILLOMAVIRUS E1 HAVING
; TITLE OF INVENTION: HELICASE ACTIVITY AND METHOD THEREFOR
; NUMBER OF SEQUENCES: 27
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/300,909
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/083,942
; FILING DATE: 01-MAY-1998
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 647 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-300-909-17

Query Match      3.3%; Score 121.5; DB 3; Length 647;
Best Local Similarity 19.7%; Pred. No. 0.0041;
Matches 110; Conservative 78; Mismatches 172; Indels 197; Gaps 26;

QY 39 GKDIGMNSYKKELOEDELK-SLQGAETTWDOSE---DMEWETT-VDEMTKKQVIFDS 92
DB 100 GKVGERT-RGTLQELSLNVSSTQATQTVYSPDGSYGMEVETAEVEVT-----149
QY 93 LVKCKLFEVLNTRKNPPGDVNWVQHEWKGDKGCHVIGGKDFPSQAQGWRRQLNYY 152
DB 150 -----VATNTNGDAEGEHGGSVREE-----CSSVDSAIDSENQDPK-----185
QY 153 WSRWLVTACNVQLTPAERIKLREIADNEWVILLTYKHKQT-----KKDYTKC 200
DB 186 -----SPTAQIKLL-LOSNNKKAAMLT-QFKETYGLSFTDLVTRFKSKDKTC 230
```



```

Db      87 YSFGIKKNPNDTVLMENMIKVPHI--KRFNSE-----LFVFKANGWQKIN-----GDEL 134
Qy      245 RHLVSKL---YTDMPETVETTTTAQETKRGRIQTKEVSIKTTLKELVHKVTSPE 301
Db      135 QGLISKMLQVLLVDYKPSL-----STLKNVVVDGLQKSTDV 169
Qy      302 WMMQPDPSYIEMMAQPGGENI-----LKNLTLEICTLT-LARTKTAFDLILEKAE--- 349
Db      170 EELVENEHYIGC-----GENNFDLNTQVVKNSIDIFPKTFLNLSLSTNDVITDKIPPYF 224
Qy      350 ---TSKLTNFSLPDTRACRIFAFHGWYVKVCHAIQCVLNRQGGKRNVTFLFHGPASTGKS 406
Db      225 KQYMLQLANYD-----DDLQVFLFOHT-AVLLTADTKYRRGLILYGGAKNGKS 271
Qy      407 IIAQAIQAQVGNVCYNAAVNFPPND-----CTNKNLIWVEAGNFGQO---VNQ 454
Db      272 VYIELVKSFF-----YSKDIVSKPLNELEGPFDESKSLIDKSLMASHEIGOSKIOEKIVND 326
Qy      455 FKAICSGQTIRIDQKGSQKIE---PTPVMITTNENITVVRIGCEBERPEHTQPIRDRML 511
Db      327 FKLLSVESMHVDRKGK--TQVEVILDKLIFSTNAILNF-----PPEHAKALERRIN 377
Qy      512 NIHLTHHL-PCDFGLVDK 528
Db      378 IIPCEYTVKADTSLIDK 395

```

Search completed: January 22, 2005, 03:43:20
 Job time : 28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:41:01 ; Search time 90.6667 Seconds
(without alignments)
2677.792 Million cell updates/sec

Title: US-10-069-056-10
Perfect score: 3637
Sequence: 1 MAGNAVSDVIGATNWLKEK.....RACFGAEPLKDPSEPLNLD 672

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

- Database :
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
 - 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	458.5	12.6	610	9	US-09-792-630-21
2	458.5	12.6	610	10	US-09-953-351-21
3	458.5	12.6	610	13	US-10-080-376-21
4	458.5	12.6	610	14	US-10-082-671-27
5	458.5	12.6	610	14	US-10-097-100-21
6	458.5	12.6	610	14	US-10-023-208-21
7	458.5	12.6	610	14	US-10-375-192-7
8	455.5	12.5	626	9	US-09-792-630-29
9	455.5	12.5	626	10	US-09-953-351-29
10	455.5	12.5	626	13	US-10-080-376-29
11	455.5	12.5	626	14	US-10-082-671-35
12	455.5	12.5	626	14	US-10-097-100-29
13	455.5	12.5	626	14	US-10-023-208-29

14	455	12.5	627	9	US-09-792-630-25	Sequence 25, Appl
15	455	12.5	627	10	US-09-953-351-25	Sequence 25, Appl
16	455	12.5	627	13	US-10-080-376-25	Sequence 25, Appl
17	455	12.5	627	14	US-10-082-671-31	Sequence 31, Appl
18	455	12.5	627	14	US-10-097-100-25	Sequence 25, Appl
19	455	12.5	627	14	US-10-023-208-25	Sequence 25, Appl
20	445.5	12.2	625	14	US-10-423-704A-3	Sequence 3, Appl
21	437.5	12.0	537	10	US-09-254-747-10	Sequence 10, Appl
22	437.5	12.0	623	9	US-09-792-630-5	Sequence 5, Appl
23	437.5	12.0	623	10	US-09-953-351-5	Sequence 5, Appl
24	437.5	12.0	623	10	US-09-254-747-2	Sequence 2, Appl
25	437.5	12.0	623	10	US-09-254-747-11	Sequence 11, Appl
26	437.5	12.0	623	13	US-10-080-376-5	Sequence 5, Appl
27	437.5	12.0	623	14	US-10-082-671-11	Sequence 11, Appl
28	437.5	12.0	623	14	US-10-097-100-5	Sequence 5, Appl
29	437.5	12.0	623	14	US-10-023-208-5	Sequence 5, Appl
30	437.5	12.0	623	14	US-10-375-192-5	Sequence 5, Appl
31	436.5	12.0	627	9	US-09-792-630-27	Sequence 27, Appl
32	436.5	12.0	627	10	US-09-953-351-27	Sequence 27, Appl
33	436.5	12.0	627	13	US-10-080-376-27	Sequence 27, Appl
34	436.5	12.0	627	14	US-10-082-671-33	Sequence 33, Appl
35	436.5	12.0	627	14	US-10-097-100-27	Sequence 27, Appl
36	436.5	12.0	627	14	US-10-023-208-27	Sequence 27, Appl
37	435.5	12.0	623	9	US-09-792-630-13	Sequence 13, Appl
38	435.5	12.0	623	10	US-09-953-351-13	Sequence 13, Appl
39	435.5	12.0	623	13	US-10-080-376-13	Sequence 13, Appl
40	435.5	12.0	623	14	US-10-082-671-19	Sequence 19, Appl
41	435.5	12.0	623	14	US-10-097-100-13	Sequence 13, Appl
42	435.5	12.0	623	14	US-10-023-208-13	Sequence 13, Appl
43	435.5	12.0	623	14	US-10-375-192-2	Sequence 2, Appl
44	431.5	11.9	546	15	US-10-696-261-7	Sequence 7, Appl
45	431.5	11.9	546	15	US-10-696-282-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1

US-09-792-630-21
; Sequence 21, Application US/09792630
; Patent No. US20020168640A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES
; FILE REFERENCE: A-70295/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/792,630
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 21
; LENGTH: 610
; TYPE: PRT
; ORGANISM: adeno-associated virus 5
US-09-792-630-21

Query Match	12.6%	Score	458.5	DB	9	Length	610		
Best Local Similarity	27.4%	Pred. No.	4.9e-31						
Matches	158	Conservative	84	Mismatches	220	Indels	115	Gaps	22
Qy	67	WQSEMEVETTVDEMTKQVFI	DSLVKKCLPEVLNTRKIF	PGDVNWFVQHWGKDQGW	126				
Db	35	WELPPESDLNLTVE--QPOLT	VADRIRRVFLYE--WNKFSQESKFFVQFKGSEY-F	88					
Qy	127	HCHVLIGGKDF-SQAQGWRRRL	QNVYWSWLTACNVOLTPAERIKLREI	-AEDNEMVT	184				
Db	89	HLHTLVETSGISSMVLGRY-	-----VSQIRAQLV----	KVVFQGLEPQINDMVA	132				
Qy	185	LLTYKHQTKQYCKVKLFGNMI	AYFLTKKISTSPPRDGGYFLSSD	SGWTKTNFLKEGE	244				
Db	133	I-----TKVKGKANKVDSGY	IPALLPK-----	VQPELOWMTNLD--	171				
Qy	245	RHLVSKLYTDDMRPETVET	TTVTVAQTKRGRITKKEVSIKTKLKS	-----LV	292				

Db 337 TNIAEAIHTVPPFGCVNNTNENFPFNDVCKMLIWWEEGKMTNKVVESAKAILGGSKVR 396
Qy 466 IDQKGSQKIEPTTPVIMTNNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGL 525
Db 397 VDQCKSSVQIDSTPVIIVTSNTNMCVVVDGNSITTFEHHQPLEDRMFKFELTKRLPPDFGK 456
Qy 526 VDKNEPIMICAWLVKNGYQSTMASYCAKWKGVDPDWE-----NWAEPKVPTPI-- 573
Db 457 ITRQEVKDFFAW-----AKVQVPVTHEFKVPRELAGTKGAESLKRPLGD 502
Qy 574 -----NLGARSAPT--TPKSTPLSONYA-LTPL 600
Db 503 VTNTSYKLEKRLARLSFVETPRSSDVTVDPAIRPL 539
RESULT 4
US-10-082-671-27
; Sequence 27, Application US/10082671
; Publication No. US20030049647A1
; GENERAL INFORMATION:
; APPLICANT: DAHIYAT, BASSIL
; APPLICANT: LI, MIN
; APPLICANT: LIU, HONG-XIANG
; TITLE OF INVENTION: USE OF NUCLEIC ACID LIBRARIES TO CREATE TOXICOLOGICAL
; TITLE OF INVENTION: PROFILES
; FILE REFERENCE: XEN/001
; CURRENT APPLICATION NUMBER: US/10/082.671
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/270,781
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 27
; LENGTH: 610
; TYPE: PRT
; ORGANISM: adeno-associated virus 5
US-10-082-671-27

Query Match 12.6%; Score 458.5; DB 14; Length 610;
Best Local Similarity 27.4%; Pred. No. 4.9e-31;
Matches 158; Conservative 84; Mismatches 220; Indels 115; Gaps 22;
Qy 67 WDQSEDEWETTVDEMTKKQVIFPDSLVKKCLFVLNKTNIFFPDGVNMFVQHEWKGQGW 126
Db 35 WELPPESDLNLTLVE--QPQLTVADRIRRVFLYE---WNKFSKQESKFFVQFEKGSY-F 88
Qy 127 HCHVLIGKDF-SQAQKQWRRQLNVYWSRWLVACNVQLTAPARIKLR-I-AEDNEWVT 184
Db 89 HLHVLTVETSGISSMWLGRY-----VSQIRALV---KVVFQIEPQINDWA 132
Qy 185 LLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPRGGYFLSSDSGKTNFLKEGE 244
Db 133 I-----TKVKKGGANKVVDGSIYPAYLLPK-----VQPELQWMTNLD-- 171
Qy 245 RHLVSKLYTDDMRPETVETTVTTAQETKRGRIQTKKEVSITKTLKE-----LV 292
Db 172 -YKLAALNLEERKRLVAQFLAESSQSEAAQ--REFSADPVKSKTSQKYMALVNLV 228
Qy 293 HKRVTSPEDEMMQPDSEYIEMMAQPGENLLKNTLEICTLTARTKTAFLILEKAETSK 352
Db 229 EHGITSKQWIOENQESYLSFNSTGNSRSQIKAALDNATKIMSLTKSAVDY----- 279
Qy 353 LTNFSLPD----TRACRIFAFHGMN--YV-KVCHAI CCVLNRQGGKNTVLFHGPASTGK 405
Db 280 LVGSSVPEDISKRIWQIFEMNGYDPAAGSILYGCW---QRSFNKNTVWLYGPATGK 336
Qy 406 SIIAQIAQAAGVNGVCYNAANVFPFNDCTNKNLIWVEEAGNFQGVNQPKAICSGQTI 465
Db 337 TNIAEAIHTVPPFGCVNNTNENFPFNDVCKMLIWWEEGKMTNKVVESAKAILGGSKVR 396
Qy 466 IDQKGSQKIEPTTPVIMTNNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGL 525
Db 397 VDQCKSSVQIDSTPVIIVTSNTNMCVVVDGNSITTFEHHQPLEDRMFKFELTKRLPPDFGK 456

Qy 526 VDKNEPIMICAWLVKNGYQSTMASYCAKWKGVDPDWE-----NWAEPKVPTPI-- 573
Db 457 ITRQEVKDFFAW-----AKVQVPVTHEFKVPRELAGTKGAESLKRPLGD 502
Qy 574 -----NLGARSAPT--TPKSTPLSONYA-LTPL 600
Db 503 VTNTSYKLEKRLARLSFVETPRSSDVTVDPAIRPL 539
RESULT 5
US-10-097-100-21
; Sequence 21, Application US/10097100
; Publication No. US20030086649A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Melander, Christian
; APPLICANT: Liu, Hong-Xiang
; APPLICANT: Jin, Cheng He
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FUSION L
; FILE REFERENCE: A-70814/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/097,100
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/953,351
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/232,960
; PRIOR FILING DATE: 2000-09-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 21
; LENGTH: 610
; TYPE: PRT
; ORGANISM: adeno-associated virus 5
US-10-097-100-21

Query Match 12.6%; Score 458.5; DB 14; Length 610;
Best Local Similarity 27.4%; Pred. No. 4.9e-31;
Matches 158; Conservative 84; Mismatches 220; Indels 115; Gaps 22;
Qy 67 WDQSEDEWETTVDEMTKKQVIFPDSLVKKCLFVLNKTNIFFPDGVNMFVQHEWKGQGW 126
Db 35 WELPPESDLNLTLVE--QPQLTVADRIRRVFLYE---WNKFSKQESKFFVQFEKGSY-F 88
Qy 127 HCHVLIGKDF-SQAQKQWRRQLNVYWSRWLVACNVQLTAPARIKLR-I-AEDNEWVT 184
Db 89 HLHVLTVETSGISSMWLGRY-----VSQIRALV---KVVFQIEPQINDWA 132
Qy 185 LLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPRGGYFLSSDSGKTNFLKEGE 244
Db 133 I-----TKVKKGGANKVVDGSIYPAYLLPK-----VQPELQWMTNLD-- 171
Qy 245 RHLVSKLYTDDMRPETVETTVTTAQETKRGRIQTKKEVSITKTLKE-----LV 292
Db 172 -YKLAALNLEERKRLVAQFLAESSQSEAAQ--REFSADPVKSKTSQKYMALVNLV 228
Qy 293 HKRVTSPEDEMMQPDSEYIEMMAQPGENLLKNTLEICTLTARTKTAFLILEKAETSK 352
Db 229 EHGITSKQWIOENQESYLSFNSTGNSRSQIKAALDNATKIMSLTKSAVDY----- 279
Qy 353 LTNFSLPD----TRACRIFAFHGMN--YV-KVCHAI CCVLNRQGGKNTVLFHGPASTGK 405
Db 280 LVGSSVPEDISKRIWQIFEMNGYDPAAGSILYGCW---QRSFNKNTVWLYGPATGK 336
Qy 406 SIIAQIAQAAGVNGVCYNAANVFPFNDCTNKNLIWVEEAGNFQGVNQPKAICSGQTI 465
Db 337 TNIAEAIHTVPPFGCVNNTNENFPFNDVCKMLIWWEEGKMTNKVVESAKAILGGSKVR 396
Qy 466 IDQKGSQKIEPTTPVIMTNNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGL 525
Db 397 VDQCKSSVQIDSTPVIIVTSNTNMCVVVDGNSITTFEHHQPLEDRMFKFELTKRLPPDFGK 456
Qy 526 VDKNEPIMICAWLVKNGYQSTMASYCAKWKGVDPDWE-----NWAEPKVPTPI-- 573

```
Db 457 ITKQEVKDFPAW-----AKVQVPTVTHEFKVPRELAGTKGAESLKRPLGD 502
Qy 574 -----NLLGSARSPT--TPKSTPLSONYA-LTFL 600
Db 503 VTNYSKLEKRLSFVPETPRSDVTVDPAPLRPL 539

RESULT 6
US-10-023-208-21
; Sequence 21, Application US/10023208
; Publication No. US20030124537A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Liu, Yuan-Ching
; TITLE OF INVENTION: PROCARYOTIC LIBRARIES AND USES
; FILE REFERENCE: A-70174-1/RT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/023,208
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 2001-12-17
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 610
; TYPE: PRT
; ORGANISM: adeno-associated virus 5
US-10-023-208-21

Query Match 12.6%; Score 458.5; DB 14; Length 610;
Best Local Similarity 27.4%; Pred. No. 4.9e-31;
Matches 158; Conservative 84; Mismatches 220; Indels 115; Gaps 22;

Qy 67 WQSEDMEWETTVDEMTHKQVFIQFDSLVKKCLFEVLNKNIFPGDVNMFVQHEWKGQGW 126
Db 35 WELPPESDLNLTVE--QPQTVADRIRRVFLYE--WNKFSKQESKFFVQFEGSEY-F 88
Qy 127 HCHVLIGGKDF-SQAQKQWRRQLNVYWSRWLTACNVQLTAPABRIKRLI-AEDNEWVT 184
Db 89 HLHVLVETSGISSMWLGRY-----VQIRALV---KVVFQIEPQINDWA 132
Qy 185 LLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPRGGYFLSSDSQWKNFLKEGE 244
Db 133 I-----TKVKKGGANKVVDGSIYPAVLLPK-----VQELQAWMTNLD-- 171
Qy 245 RHLVSKLYTDDMPETVETTTAQTGRGRIQTKKEVSIKTLKE-----LV 292
Db 172 -YKLAALNLEERKRLVAQFLAESSQSQBAASQ--REFSADPVIKSKTSOKYVALVNWLV 228
Qy 293 HKRVTSPEMMQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSK 352
Db 229 EHGITSQKQIENQESYLSFNSSTGNSRSQIKAAALDNATKIMSLTKSAVDY----- 279
Qy 353 LTNFSLPD----TRACRIPAFHGN--YV-KVCHAI CCVNLNRQGGKNTVLFHGPASTGK 405
Db 280 LVGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGCW---QRSFNKNTVWLYGPATTKG 336
Qy 406 SIATAQAQAVGNVGCYNAANVPFNDCTNKNLIWVEEAGNFQGVQVQKAIKCSGOTIR 465
Db 337 TNIAEALAHVTPFGCVNWTNENFPFNDCTNKNLIWVEEAGNTKNVVEAKAILGSKVR 396
Qy 466 IDQKGGSKQIEPTPVIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDGL 525
Db 397 VDQCKSSVQIDSTPVIIVTSNTNMCVVVDGNSITTFHQQLPDRMPKFEFTKELPDPFGK 456
Qy 526 VDKNEWPMICANLVKNGYQSTMASYCAKWKGVDPDWE-----NWAEPKVPPI-- 573
Db 457 ITKQEVKDFPAW-----AKVQVPTVTHEFKVPRELAGTKGAESLKRPLGD 502

Qy 574 -----NLLGSARSPT--TPKSTPLSONYA-LTFL 600
Db 503 VTNYSKLEKRLSFVPETPRSDVTVDPAPLRPL 539
```

RESULT 8

```
US-09-792-630-29
; Sequence 29, Application US/09792630
; Patent No. US20020168640A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
```

```
RESULT 7
US-10-375-192-7
; Sequence 7, Application US/10375192
; Publication No. US20030224404A1
; GENERAL INFORMATION:
; APPLICANT: Vega, Manuel
; APPLICANT: Drittanti, Lila
; TITLE OF INVENTION: HIGH-THROUGHPUT DIRECTED EVOLUTION OF NUCLEIC ACIDS BY RATIONAL
; FILE REFERENCE: 37851-918
; CURRENT APPLICATION NUMBER: US/10/375,192
; CURRENT FILING DATE: 2003-02-24
; PRIOR APPLICATION NUMBER: US 60/360,085
; PRIOR FILING DATE: 2002-02-25
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Adeno-associated virus 5
US-10-375-192-7

Query Match 12.6%; Score 458.5; DB 14; Length 610;
Best Local Similarity 27.4%; Pred. No. 4.9e-31;
Matches 158; Conservative 84; Mismatches 220; Indels 115; Gaps 22;

Qy 67 WQSEDMEWETTVDEMTHKQVFIQFDSLVKKCLFEVLNKNIFPGDVNMFVQHEWKGQGW 126
Db 35 WELPPESDLNLTVE--QPQTVADRIRRVFLYE--WNKFSKQESKFFVQFEGSEY-F 88
Qy 127 HCHVLIGGKDF-SQAQKQWRRQLNVYWSRWLTACNVQLTAPABRIKRLI-AEDNEWVT 184
Db 89 HLHVLVETSGISSMWLGRY-----VQIRALV---KVVFQIEPQINDWA 132
Qy 185 LLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPRGGYFLSSDSQWKNFLKEGE 244
Db 133 I-----TKVKKGGANKVVDGSIYPAVLLPK-----VQELQAWMTNLD-- 171
Qy 245 RHLVSKLYTDDMPETVETTTAQTGRGRIQTKKEVSIKTLKE-----LV 292
Db 172 -YKLAALNLEERKRLVAQFLAESSQSQBAASQ--REFSADPVIKSKTSOKYVALVNWLV 228
Qy 293 HKRVTSPEMMQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSK 352
Db 229 EHGITSQKQIENQESYLSFNSSTGNSRSQIKAAALDNATKIMSLTKSAVDY----- 279
Qy 353 LTNFSLPD----TRACRIPAFHGN--YV-KVCHAI CCVNLNRQGGKNTVLFHGPASTGK 405
Db 280 LVGSSVPEDISKRIWQIFEMNGYDPAYAGSILYGCW---QRSFNKNTVWLYGPATTKG 336
Qy 406 SIATAQAQAVGNVGCYNAANVPFNDCTNKNLIWVEEAGNFQGVQVQKAIKCSGOTIR 465
Db 337 TNIAEALAHVTPFGCVNWTNENFPFNDCTNKNLIWVEEAGNTKNVVEAKAILGSKVR 396
Qy 466 IDQKGGSKQIEPTPVIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDGL 525
Db 397 VDQCKSSVQIDSTPVIIVTSNTNMCVVVDGNSITTFHQQLPDRMPKFEFTKELPDPFGK 456
Qy 526 VDKNEWPMICANLVKNGYQSTMASYCAKWKGVDPDWE-----NWAEPKVPPI-- 573
Db 457 ITKQEVKDFPAW-----AKVQVPTVTHEFKVPRELAGTKGAESLKRPLGD 502

Qy 574 -----NLLGSARSPT--TPKSTPLSONYA-LTFL 600
Db 503 VTNYSKLEKRLSFVPETPRSDVTVDPAPLRPL 539
```

; APPLICANT: Dahiyat, Basail I.
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES
; FILE REFERENCE: A-70295/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/792.630
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Muscovy duck parvovirus
US-09-792-630-29

Query Match 12.5%; Score 455.5; DB 9; Length 626;
Best Local Similarity 25.2%; Pred. No. 9.3e-31;
Matches 164; Conservative 82; Mismatches 264; Indels 141; Gaps 22;
QY 75 WEYTT---VDEMTKKQVFIFDSLVKKCLFEVLNTKNIF-----PGDWNWVQHEWGK 122
DB 45 WEPTGIWNEHVNLPMWTLADKI-----KNIFQRMNQFNQDETDFFQLEEGS 93
QY 123 DOGWCHVLGGKDFSOAQGKWMRRQLNVYWSRWLVACNVQLTPAERIK---LREIAED 179
DB 94 EY-IHLHAVCPGCRSFVLGRY-----MSQIKDSILRDVYEG 129
QY 180 N-----EWVTLTYKHQTKDYTKCVLFGNMIAYYFLTKKISTSPRRGGYFLSDSG 234
DB 130 KQVKIPDWFSTITKRGQGNKVTAA-----YILHYLIPKKQ-----PELQWAFNTMPLF 179
QY 235 WKTNFKEGEHLVSKLYTDDMPETVETVTTTAQETKRGRIQTKKEVSIKTKLKVHK 294
DB 180 TAAALCQKROELDLAFQSEMNNAVQEOASTAPL-----ISNRAKNYSNLVDWLIEM 235
QY 295 RVTSPEDDMMQPDYSYEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSKLT 354
DB 236 GITSEKQWLTENKESYRSFQATSSNNRQVKAALENARAEMLLTKTATDYLGKDPVLDT 295
QY 355 NFSLPDTRACRIFAFHGWNVYKCHACCVLNRQGGKRNVTLPFGPASTGKSIQAQIAQ 414
DB 296 K-----NRIYQILKLNYPQYVGVSLCGWVKREFNKNRINAIWLYGPATTKTNIAEIAH 350
QY 415 AVGNVCYNAANVFPENDCTKNLIWVEAGNFGQOVNPFKALCSQOTIRIDOKGKSK 474
DB 351 AVFYGCNVNTNENFPFNDCCVDKMLIWEEGKMTNKVVESAKAILGGSVRVDQCKGVS 410
QY 475 QIEPTVIMTTNENITVVRIGCSERPEHTQPIRDRMLNIHLTHLPDGLVDKNE--- 530
DB 411 CIEPTVLIISNTDMCMVDGNSTTMEHRIPLEERMFQIVLSHKGNGFKISKKEVKEF 470
QY 531 --W-----PMICAWLVKNGYQSTMASYCAKWKGVDPWSN-----WAEPK----- 568
DB 471 FKWANDNLVPVWSEFKVPTNEQTKLTE-----PVPERANEPEPPKIWAPPTREELEE 524
QY 569 -----VPTPINLGSARSPTTKSTPLSQNYALTPLASDLEDLALPESWTPNT 617
DB 525 LRASPELFASVAPLP-----SSPDTSPKRRKTRGEYQVRCAMHSL-DNSMNVFECLEC 576
QY 618 PVAGTAETQNTGAGSKAC-QDGQLSPTWSEIEEDLRACFGAPLKKDFSE 667
DB 577 ERANFPFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 612

RESULT 9
US-09-953-351-29
; Sequence 29, Application US/09953351
; Publication No. US20030036643A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Melander, Christian
; APPLICANT: Liu, Hong-Xiang
; APPLICANT: Jin, Cheng He
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FUSION
; FILE REFERENCE: A-70814/RFT/RMS/RMK

; CURRENT APPLICATION NUMBER: US/09/953.351
; CURRENT FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/232,960
; PRIOR FILING DATE: 2000-09-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Muscovy duck parvovirus
US-09-953-351-29

Query Match 12.5%; Score 455.5; DB 10; Length 626;
Best Local Similarity 25.2%; Pred. No. 9.3e-31;
Matches 164; Conservative 82; Mismatches 264; Indels 141; Gaps 22;
QY 75 WEYTT---VDEMTKKQVFIFDSLVKKCLFEVLNTKNIF-----PGDWNWVQHEWGK 122
DB 45 WEPTGIWNEHVNLPMWTLADKI-----KNIFQRMNQFNQDETDFFQLEEGS 93
QY 123 DOGWCHVLGGKDFSOAQGKWMRRQLNVYWSRWLVACNVQLTPAERIK---LREIAED 179
DB 94 EY-IHLHAVCPGCRSFVLGRY-----MSQIKDSILRDVYEG 129
QY 180 N-----EWVTLTYKHQTKDYTKCVLFGNMIAYYFLTKKISTSPRRGGYFLSDSG 234
DB 130 KQVKIPDWFSTITKRGQGNKVTAA-----YILHYLIPKKQ-----PELQWAFNTMPLF 179
QY 235 WKTNFKEGEHLVSKLYTDDMPETVETVTTTAQETKRGRIQTKKEVSIKTKLKVHK 294
DB 180 TAAALCQKROELDLAFQSEMNNAVQEOASTAPL-----ISNRAKNYSNLVDWLIEM 235
QY 295 RVTSPEDDMMQPDYSYEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSKLT 354
DB 236 GITSEKQWLTENKESYRSFQATSSNNRQVKAALENARAEMLLTKTATDYLGKDPVLDT 295
QY 355 NFSLPDTRACRIFAFHGWNVYKCHACCVLNRQGGKRNVTLPFGPASTGKSIQAQIAQ 414
DB 296 K-----NRIYQILKLNYPQYVGVSLCGWVKREFNKNRINAIWLYGPATTKTNIAEIAH 350
QY 415 AVGNVCYNAANVFPENDCTKNLIWVEAGNFGQOVNPFKALCSQOTIRIDOKGKSK 474
DB 351 AVFYGCNVNTNENFPFNDCCVDKMLIWEEGKMTNKVVESAKAILGGSVRVDQCKGVS 410
QY 475 QIEPTVIMTTNENITVVRIGCSERPEHTQPIRDRMLNIHLTHLPDGLVDKNE--- 530
DB 411 CIEPTVLIISNTDMCMVDGNSTTMEHRIPLEERMFQIVLSHKGNGFKISKKEVKEF 470
QY 531 --W-----PMICAWLVKNGYQSTMASYCAKWKGVDPWSN-----WAEPK----- 568
DB 471 FKWANDNLVPVWSEFKVPTNEQTKLTE-----PVPERANEPEPPKIWAPPTREELEE 524
QY 569 -----VPTPINLGSARSPTTKSTPLSQNYALTPLASDLEDLALPESWTPNT 617
DB 525 LRASPELFASVAPLP-----SSPDTSPKRRKTRGEYQVRCAMHSL-DNSMNVFECLEC 576
QY 618 PVAGTAETQNTGAGSKAC-QDGQLSPTWSEIEEDLRACFGAPLKKDFSE 667
DB 577 ERANFPFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 612

RESULT 10
US-10-080-376-29
; Sequence 29, Application US/10080376
; Publication No. US20020172968A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Dahiyat, Basail I.
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES
; FILE REFERENCE: A-70295-2/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/080.376
; CURRENT FILING DATE: 2000-02-19
; PRIOR APPLICATION NUMBER: US 09/792,630

; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Muscovy duck parvovirus
US-10-080-376-29

Query Match 12.5%; Score 455.5; DB 13; Length 626;
Best Local Similarity 25.2%; Pred. No. 9.3e-31;
Matches 164; Conservative 82; Mismatches 264; Indels 141; Gaps 22;
Qy 75 WETT-----VDEMTKKQVIFDLSLVKCLFVLTNKNIP-----PGDVNMFVQHEWGK 122
Db 45 WEPTGIWNMEHVNLPMTVLADKI-----KNIFIQRWNPQNODETDFPQLEEGS 93
Qy 123 DOGWCHVLIGGKDFSOAQGWRRQLNVVWSRWLTACNVQLTPAERIK---LREIAED 179
Db 94 EY-IHLHVAVCPGECRSFVLGRY-----MSQIKDSILRDVVEG 129
Qy 180 N-----EWVTLTYKHQTKDYTKCVLFGNMIAVYFLTKKISTSPRDGGYFLSSDSG 234
Db 130 KQVKIPDWSITTKRGQNKVTAA-----YILHYLIPKQ-----PELQWAFNNPLF 179
Qy 235 WKTNFKEGEHLVSKLYTDDMRPETVETVTTTAQETKRGRIQTKKEVSITKTLKELVHK 294
Db 180 TAAALCLOKQELLDAPQESNNVAVQEDQASTAAPL-----ISNRAKNVSNLVDWLIEM 235
Qy 295 RVTSPEMMQPPSYIEMMAQPGENLLKNTLEICTLTARTKTAFDLILEKAETSCLT 354
Db 236 GITSEKQWLTKENKESYRSFQATSSNNRQVKAALENARAEMLLTKTATDYLIGKDPVLDIT 295
Qy 355 NFSLPDTFRACIFAFHGNVYKVKCHAICCVLRQGGKRNVLPHGPASTGKSIIAQIAQ 414
Db 296 K-----NRIYQILKLNYPQVGSVLCGWVRFKRNAINWLYGPATTKTNIAEIAH 350
Qy 415 AVGVGVCYNAANVFPNDCTNKNLIWVEEAGNFGQVQNFKAICSGQTRIDQKGGSK 474
Db 351 AVPFVGVNWTNENFPNDCTNKNLIWVEEAGNFGQVQNFKAICSGQTRIDQKGGSK 410
Qy 475 QIETPTVIMTNNITVVRIGCEERDEHTQPIRDMNLNHLTHLPGDGLVDKNE-----530
Db 411 CIEPTVITNTDMCMIVDGNSTTMEHRIPLEERMFQIVLSHKLGNFGKISKEVKEF 470
Qy 531 --W-----PMICAWLVKNGYQSTMAASYCAKWKGVDPDSEN-----WAEPK-----568
Db 471 FKWANDNLVPVSEFKVPTNEQTKLTE-----PVPERANESEPPKIWAPPTREELEEI 524
Qy 569 -----VPTPINLLGSARSPTTPKSTPLSONYALTPLASDLIEDLALPWSPTNT 617
Db 525 LRASPELFASVAPLP-----SSPDTSPKPKTKRGEYQVRCAMHSL-DNSMNVFECLEC 576
Qy 618 PVAGTATONTGEAGSKAC-QDGLSPTWSEIEEDLRACFGAEPLKDFSE 667
Db 577 ERANFPFQSLGE---NFCNQHG-----WYD-----CAFCNELKODMNE 612

RESULT 11

US-10-082-671-35
; Sequence 35, Application US/10082671
; Publication No. US20030049647A1
; GENERAL INFORMATION:
; APPLICANT: DAHIYAT, BASSIL
; APPLICANT: LI, MIN
; TITLE OF INVENTION: USE OF NUCLEIC ACID LIBRARIES TO CREATE TOXICOLOGICAL
; FILE OF INVENTION: PROFILES
; FILE REFERENCE: XEN/001
; CURRENT APPLICATION NUMBER: US/10/082,671
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/270,781
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 58

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Muscovy duck parvovirus
US-10-082-671-35

Query Match 12.5%; Score 455.5; DB 14; Length 626;
Best Local Similarity 25.2%; Pred. No. 9.3e-31;
Matches 164; Conservative 82; Mismatches 264; Indels 141; Gaps 22;
Qy 75 WETT-----VDEMTKKQVIFDLSLVKCLFVLTNKNIP-----PGDVNMFVQHEWGK 122
Db 45 WEPTGIWNMEHVNLPMTVLADKI-----KNIFIQRWNPQNODETDFPQLEEGS 93
Qy 123 DOGWCHVLIGGKDFSOAQGWRRQLNVVWSRWLTACNVQLTPAERIK---LREIAED 179
Db 94 EY-IHLHVAVCPGECRSFVLGRY-----MSQIKDSILRDVVEG 129
Qy 180 N-----EWVTLTYKHQTKDYTKCVLFGNMIAVYFLTKKISTSPRDGGYFLSSDSG 234
Db 130 KQVKIPDWSITTKRGQNKVTAA-----YILHYLIPKQ-----PELQWAFNNPLF 179
Qy 235 WKTNFKEGEHLVSKLYTDDMRPETVETVTTTAQETKRGRIQTKKEVSITKTLKELVHK 294
Db 180 TAAALCLOKQELLDAPQESNNVAVQEDQASTAAPL-----ISNRAKNVSNLVDWLIEM 235
Qy 295 RVTSPEMMQPPSYIEMMAQPGENLLKNTLEICTLTARTKTAFDLILEKAETSCLT 354
Db 236 GITSEKQWLTKENKESYRSFQATSSNNRQVKAALENARAEMLLTKTATDYLIGKDPVLDIT 295
Qy 355 NFSLPDTFRACIFAFHGNVYKVKCHAICCVLRQGGKRNVLPHGPASTGKSIIAQIAQ 414
Db 296 K-----NRIYQILKLNYPQVGSVLCGWVRFKRNAINWLYGPATTKTNIAEIAH 350
Qy 415 AVGVGVCYNAANVFPNDCTNKNLIWVEEAGNFGQVQNFKAICSGQTRIDQKGGSK 474
Db 351 AVPFVGVNWTNENFPNDCTNKNLIWVEEAGNFGQVQNFKAICSGQTRIDQKGGSK 410
Qy 475 QIETPTVIMTNNITVVRIGCEERDEHTQPIRDMNLNHLTHLPGDGLVDKNE-----530
Db 411 CIEPTVITNTDMCMIVDGNSTTMEHRIPLEERMFQIVLSHKLGNFGKISKEVKEF 470
Qy 531 --W-----PMICAWLVKNGYQSTMAASYCAKWKGVDPDSEN-----WAEPK-----568
Db 471 FKWANDNLVPVSEFKVPTNEQTKLTE-----PVPERANESEPPKIWAPPTREELEEI 524
Qy 569 -----VPTPINLLGSARSPTTPKSTPLSONYALTPLASDLIEDLALPWSPTNT 617
Db 525 LRASPELFASVAPLP-----SSPDTSPKPKTKRGEYQVRCAMHSL-DNSMNVFECLEC 576
Qy 618 PVAGTATONTGEAGSKAC-QDGLSPTWSEIEEDLRACFGAEPLKDFSE 667
Db 577 ERANFPFQSLGE---NFCNQHG-----WYD-----CAFCNELKODMNE 612

RESULT 12

US-10-097-100-29
; Sequence 29, Application US/10097100
; Publication No. US20030068649A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Melander, Christian
; APPLICANT: Liu, Hong-Xiang
; APPLICANT: Jin, Cheng He
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FUSION LI
; FILE REFERENCE: A-70814/RET/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/097,100
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/953,351
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/232,960
; PRIOR FILING DATE: 2000-09-14

; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Muscovy duck parvovirus
US-10-097-100-29

Query Match 12.5%; Score 455.5; DB 14; Length 626;
Best Local Similarity 25.2%; Pred. No. 9.3e-31;
Matches 164; Conservative 82; Mismatches 264; Indels 141; Gaps 22;
Qy 75 WETT-----VDEMTKKQVFIPDSLVKKCLFEVLNTKNIF-----PGDVNWFVQHEWGK 122
Db 45 WEPTGIWMEHVNLPMTLADKI-----KNIFIQRWNQFNQDETFDFQLEEGS 93
Qy 123 DQGWCHVLIGGKDFSOAQGWRRQLNVVWSRWLTACNVQLTTPAERIK---LREIAED 179
Db 94 EY-IHLHVAVCPGCRSFVLGRY-----MSQIKDSILRDVYEG 129
Qy 180 N-----EWVTLTYKHKQTKDYTKCVLFGNMIAYYFLTKKISTSPPRDGGYFLSSDSG 234
Db 130 KQVKIPDMFSITTKRGQNKVTAA-----YILHYLIPKKQ-----PELQWAFNMPLF 179
Qy 235 WKTNFKGEBRHLVSKLYTDDMRPETVETVTTTAQSTKRGRIQTKKEVSIKTLKELVHK 294
Db 180 TAAALCLQKQELLDAFQSEMNNAVQEDQASTAAPL-----ISNRAKNYSNLVDMLIEM 235
Qy 295 RVTSPEMMQPDYSIEMMAQPGENLLKNTLEICTLTARTKTAFDLILEKAETSKLT 354
Db 236 GITSEKQWLTENKESYRSFQATSSNNRQVKAALENARAEMLLTKTATDYLGKDPVLDT 295
Qy 355 NFSLPDTRACRIFAFHGWNVYKCHACCVLNRQGGKRNVTLPFGPASTGKSIIAQIAIAQ 414
Db 236 GITSEKQWLTENKESYRSFQATSSNNRQVKAALENARAEMLLTKTATDYLGKDPVLDT 295
Qy 355 NFSLPDTRACRIFAFHGWNVYKCHACCVLNRQGGKRNVTLPFGPASTGKSIIAQIAIAQ 414
Db 296 K-----NRIYQILKLNYPQYVGSVLCGWKREFNKNAINWLYGPATTKGTNIAEIAH 350
Qy 415 AVGVGVCYNAANVPFNDCTNKNLIWBEAGNFGQOVNOFKAICSGQTTIRIOKKGSK 474
Db 351 AVDFYGCYVNTNENFPDNCVDRKMLIWEKGKTNKVESAKAILGSAVRVDQCKGVS 410
Qy 475 QIBPTPVMITNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNE--- 530
Db 411 CIBPTPVIITNTDMCMIVDGNSTTMEHRIPLEERMFQIVLSHKLGNFGKISKEVKEF 470
Qy 531 --W-----PMICAWLVKNGYQSTMASYCAKWKGVDPWSEN-----WAEPK----- 568
Db 471 FKWANDLVVWSEFKVPTNEQTKLTE-----PVPERANESEPPKIWAPPTREELEEI 524
Qy 569 -----VPTPINLLGSARSPTTPKSTPLSQNYALTPLASDLEDALEPWSPTNT 617
Db 525 LRASPELFASVAPLP-----SSPDTSPKRRKTRGEYQVRCAMHSL-DNSMNVFECLEC 576
Qy 618 PVAGTAETQNTGEAGSKAC-QDQGLSPTWSEIEEDLRACFGABPLKDFSE 667
Db 577 ERANFPFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 612

RESULT 13
US-10-023-208-29
; Sequence 29, Application US/10023208
; Publication No. US20030124537A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Liu, Yuan-Ching
; TITLE OF INVENTION: PROCARYOTIC LIBRARIES AND USES
; FILE REFERENCE: A-70174-1/RT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/023,208
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/256,163
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29

; LENGTH: 626
; TYPE: PRT
; ORGANISM: Muscovy duck parvovirus
US-10-023-208-29

Query Match 12.5%; Score 455.5; DB 14; Length 626;
Best Local Similarity 25.2%; Pred. No. 9.3e-31;
Matches 164; Conservative 82; Mismatches 264; Indels 141; Gaps 22;
Qy 75 WETT-----VDEMTKKQVFIPDSLVKKCLFEVLNTKNIF-----PGDVNWFVQHEWGK 122
Db 45 WEPTGIWMEHVNLPMTLADKI-----KNIFIQRWNQFNQDETFDFQLEEGS 93
Qy 123 DQGWCHVLIGGKDFSOAQGWRRQLNVVWSRWLTACNVQLTTPAERIK---LREIAED 179
Db 94 EY-IHLHVAVCPGCRSFVLGRY-----MSQIKDSILRDVYEG 129
Qy 180 N-----EWVTLTYKHKQTKDYTKCVLFGNMIAYYFLTKKISTSPPRDGGYFLSSDSG 234
Db 130 KQVKIPDMFSITTKRGQNKVTAA-----YILHYLIPKKQ-----PELQWAFNMPLF 179
Qy 235 WKTNFKGEBRHLVSKLYTDDMRPETVETVTTTAQSTKRGRIQTKKEVSIKTLKELVHK 294
Db 180 TAAALCLQKQELLDAFQSEMNNAVQEDQASTAAPL-----ISNRAKNYSNLVDMLIEM 235
Qy 295 RVTSPEMMQPDYSIEMMAQPGENLLKNTLEICTLTARTKTAFDLILEKAETSKLT 354
Db 236 GITSEKQWLTENKESYRSFQATSSNNRQVKAALENARAEMLLTKTATDYLGKDPVLDT 295
Qy 355 NFSLPDTRACRIFAFHGWNVYKCHACCVLNRQGGKRNVTLPFGPASTGKSIIAQIAIAQ 414
Db 296 K-----NRIYQILKLNYPQYVGSVLCGWKREFNKNAINWLYGPATTKGTNIAEIAH 350
Qy 415 AVGVGVCYNAANVPFNDCTNKNLIWBEAGNFGQOVNOFKAICSGQTTIRIOKKGSK 474
Db 351 AVDFYGCYVNTNENFPDNCVDRKMLIWEKGKTNKVESAKAILGSAVRVDQCKGVS 410
Qy 475 QIBPTPVMITNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNE--- 530
Db 411 CIBPTPVIITNTDMCMIVDGNSTTMEHRIPLEERMFQIVLSHKLGNFGKISKEVKEF 470
Qy 531 --W-----PMICAWLVKNGYQSTMASYCAKWKGVDPWSEN-----WAEPK----- 568
Db 471 FKWANDLVVWSEFKVPTNEQTKLTE-----PVPERANESEPPKIWAPPTREELEEI 524
Qy 569 -----VPTPINLLGSARSPTTPKSTPLSQNYALTPLASDLEDALEPWSPTNT 617
Db 525 LRASPELFASVAPLP-----SSPDTSPKRRKTRGEYQVRCAMHSL-DNSMNVFECLEC 576
Qy 618 PVAGTAETQNTGEAGSKAC-QDQGLSPTWSEIEEDLRACFGABPLKDFSE 667
Db 577 ERANFPFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 612

RESULT 14
US-09-792-630-25
; Sequence 25, Application US/09792630
; Patent No. US20020168640A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Dahiayat, Bassil I.
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES
; FILE REFERENCE: A-70295/RT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/792,630
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Barbarie duck parvovirus
US-09-792-630-25

Best Local Similarity	25.8%;	Pred. No. 1e-30;	
Matches 168;	Conservative	76;	Mismatches 265; Indels 140; Gaps 23;
QY	75	WETT-----VDEMTKQKQVIFDSLVKKCLFEVLNTKNIF-----PGDVNWFVQHEWGK	122
Db	45	WEPTGIWNMEHVNLPWMTLADK-----KNIFIQRMNQFNQDEDTFFPQEEGS	93
QY	123	DOGHCHVLIGGKDFSAQOGKWRRLQNLVYWSRWLVTAACNVQLTPAERIK---LREIAED	179
Db	94	EY-IHLHCCI-----AQNVRSEFVLGRYS-----QIKDSLRLDVYEG	130
QY	180	N-----BWTLLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPRPGGVFLSSDSG	234
Db	131	KQVKIPDWFISITTKRGQGNKVTAA-----YILHVLIPKKQ-----PELQWAFNTNPLF	180
QY	235	WKTNFLKEGRHVLKSLYTDMPETVETVTVTAQETKGRIOQTKKEVSIKTTLKELVHK	294
Db	181	TAAALCQKQKQELLDAPQSEEMNAVQEDQSAAPL-----ISNRAKNYSNVLWMIEM	236
QY	295	RVTSPEDDMMQPDSDYIEMMAQPGGEGNLLKNTLIEICTLTLARTKTAFDLILEKAETSKLT	354
Db	237	GITSEKQWLTKENKESYRSFQATSSNNRQVKAALENARAEMLLTKTATDYLGKDPVLDDIT	296
QY	355	NFSLPDTTRACRIIFAPHGWNVTKVCHALCCVLNRQGGKRNVTLVPHGPASTKGSIIAQIAIQ	414
Db	297	K-----NRIYQILKLANNYNPQYVGSVLCGWKVRKFNKRNAILWYLGPAITTKTNIAEAIAH	351
QY	415	AVGNVGCYNAANVNFPPDCTNKNLVWEAGNFGQVNFQKAI CSGOTIRIDOKGKSGK	474
Db	352	AVFPYGCVWNTEFNFFDNCVDKMLLWEEGQMTNKVVSAKAILGSAVRVDOKCKGSV	411
QY	475	QIEPTPVIMTNNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLHPDFGLVDKNE----	530
Db	412	CIEPTVIIISNTDMCMIVDGNSTTMEHRIPLERMFQIIVLSHKLGNFGKISKEVKEF	471
QY	531	--W-----PMICAWLVKNGYQSTMASYCAKWKGVPDWSEN-----WAEPK-----	568
Db	472	FKWANDNLVPEVWSEFKVPTNEQTKLTE-----PVPERANEPSEPPKIIWAPPTRESLEBI	525
QY	569	-----VPTPINLLGSARSPTTPKTPLSQNYALTPLASDLELDLALPEWTPNT	617
Db	526	LRASPELFASVAPLP-----SSPDTSPKPKTRGEYQVRCAMHSL--DNSMNVFECLEC	577
QY	618	PVAGTAEQTQNTGAGSKAC-QDQQLSGPTWSEIIEEDLRACFGAEPLKDPFSE	667
Db	578	ERANFPFESFOSLGS--NFCNQHG-----WYD-----CAFCNELKDDWNE	613

Search completed: January 22, 2005, 04:03:42
Job time : 92.6667 secs

Query Match 12.5%; Score 455; DB 10; Length 627;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:27:18 ; Search time 22.3333 Seconds
(without alignments)
2895.119 Million cell updates/sec

Title: us-10-069-056-10
Perfect score: 3637
Sequence: 1 MAGNAYSDEVLGATNWLKSK.....RACFGAEPLKKDPSEPLNLD 672

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3623	99.6	672	1 UYPVIM	noncapsid protein
2	3556	97.8	721	1 UYPVIM	noncapsid protein
3	3326	91.4	672	1 UYPV1	noncapsid protein
4	3322	91.3	668	1 A44276	noncapsid protein
5	2647	72.8	668	1 UYPVME	noncapsid protein
6	2644	72.7	668	1 UYPVPP	noncapsid protein
7	2642	72.6	668	1 UYPVCP	noncapsid protein
8	2485	68.3	660	1 UYPVPP	noncapsid protein
9	2476	68.1	662	1 UYPVNA	noncapsid protein
10	1621.5	44.6	392	1 UYPVIF	noncapsid protein
11	550	15.1	641	2 S41439	gene NS-1 protein
12	544.5	15.0	641	2 S41434	gene NS-1 protein
13	543	14.9	620	1 UYPVAP	noncapsid protein
14	533.5	14.7	641	2 S41861	gene NS-1 protein
15	455.5	12.5	626	2 S52209	noncapsid protein
16	424.5	11.7	536	1 UYADIA	noncapsid protein
17	352	9.7	726	1 UYPV51	noncapsid protein
18	348	9.6	671	1 UYPV19	noncapsid protein
19	169	4.6	490	2 T44050	hypothetical prote
20	168.5	4.6	490	2 JQ1630	noncapsid protein
21	144	4.0	276	2 S26428	hypothetical 31.5K
22	131.5	3.6	849	1 UYPVAD	noncapsid protein
23	127	3.5	545	2 B44054	orf2 protein - Jun
24	121.5	3.3	614	1 W1W141	E1 protein - human
25	121.5	3.3	647	1 W1W139	E1 protein - human
26	120.5	3.3	825	2 G96665	protein F22C12.12
27	117.5	3.2	1008	2 T18832	probable RNA helic
28	114.5	3.1	1098	2 B70232	hypothetical prote
29	113.5	3.1	497	2 C70454	transcription regu

RESULT 1

UYPVIM

noncapsid protein NS1 - minute virus of mice
C;Species: minute virus of mice, murine parvovirus
C;Date: 14-Nov-1983 #sequence_revision 28-Aug-1985 #text_change 09-Jul-2004
C;Accession: A03696
R;Astell, C.R.; Thomson, M.; Merchlinsky, M.; Ward, D.C.
Nucleic Acids Res. 11, 999-1018, 1983

A;Title: The complete DNA sequence of minute virus of mice, an autonomous parvovirus.
A;Reference number: A03696; MUID:83143341; PMID:6298737
A;Accession: A03696
A;Molecule type: DNA
A;Residues: 1-672 <AST>

A;Cross-references: UNIPROT:P03134; EMBL:V0115

A;Superfamily: parvovirus noncapsid protein

C;Keywords: noncapsid protein

Query Match 99.6%; Score 3623; DB 1; Length 672;
Best Local Similarity 99.7%; Pred. No. 9.2e-255;
Matches 670; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	1	MAGNAYSDEVLGATNWLKESKNOEVSFVFNENVQLNGKDIGNSYKKLODELKSLQ	60
Db	1	MAGNAYSDEVLGATNWLKESKNOEVSFVFNENVQLNGKDIGNSYKKLODELKSLQ	60
Qy	61	RGAEETWDQSEDMEWETTVDKTKQVFIQVDSLVKKCLFEVLNTKNIIFPGDVNMFVQHEW	120
Db	61	RGAEETWDQSEDMEWETTVDKTKQVFIQVDSLVKKCLFEVLNTKNIIFPGDVNMFVQHEW	120
Qy	121	GKQGWCHVLIIGKDFSQAGKWRRLQNLVYWSRWLVTAACNVQLTPAERIKLREIAEDN	180
Db	121	GKQGWCHVLIIGKDFSQAGKWRRLQNLVYWSRWLVTAACNVQLTPAERIKLREIAEDN	180
Qy	181	EWTLTYKHQKQKDYTKCVLFGNMIAYYFLTKKISTSPRDGGVFLSDSGWKTNFL	240
Db	181	EWTLTYKHQKQKDYTKCVLFGNMIAYYFLTKKISTSPRDGGVFLSDSGWKTNFL	240
Qy	241	KEGERHLVSKLYTDDMRPETVETTTTAQSTKRGRIQTKKEVSIKTTLKELVHKRVTSPE	300
Db	241	KEGERHLVSKLYTDDMRPETVETTTTAQSTKRGRIQTKKEVSIKTTLKELVHKRVTSPE	300
Qy	301	DNMMPDSDSIEMMAQPGGNNLKNLLEICTLTARTKTAFLDILEKAEKSKLTNFSPLD	360
Db	301	DNMMPDSDSIEMMAQPGGNNLKNLLEICTLTARTKTAFLDILEKAEKSKLTNFSPLD	360
Qy	361	TRACRIFAFHGWYVVKVCHACVILNRQGGKRVTLFPHGPASTGKSIIAQAIQAVGNVG	420
Db	361	TRTCRIFAFHGWYVVKVCHACVILNRQGGKRVTLFPHGPASTGKSIIAQAIQAVGNVG	420
Qy	421	CYNAANVNFNDCTNKNLIWVEAGNFQGVNQFKAICSGQTRIDQKGGKQIETPTP	480
Db	421	CYNAANVNFNDCTNKNLIWVEAGNFQGVNQFKAICSGQTRIDQKGGKQIETPTP	480

E1 protein - human
E1 protein - human
anfA protein - Azo
hypothetical prote
E1 protein - human
probable ATP bindi
structural polypro
hypothetical prote
hypothetical prote
E1 protein - human
proteinase ClpX [i
dystrophin - mouse
hypothetical WW do
type I site-specif
TipC protein - all
large T antigen -


```
QY 421 CYNAAVNFPPNDCTNKNLIWVEEAGNFGQOVNQFKAICSGQIRIDQKKGSKQIEPTP 480
DB 421 CYNAAVNFPPNDCTNKNLIWVEEAGNFGQOVNQFKAICSGQIRIDQKKGSKQIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNWPMICAWLVK 540
DB 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNWPMICAWLVK 540
QY 541 NGYSTMASYCAKWKGVDPDSENWAEKPYPTPINLLGSARSPTTTPKSTPLSQNYALTPL 600
DB 541 NGYSTMASYCAKWKGVDPDSENWAEKPYPTPINLLGSARSPTTTPKSTPLSQNYALTPL 600
QY 601 ASLEDLALEPWPSTPNTPVAGTAEQTONTGSAKQACODGQLSPTWSEIEEDLRACFGAEP 660
DB 601 ASLEDLALEPWPSTPNTPVAGTAEQTONTGSAKQACODGQLSPTWSEIEEDLRACFGAEP 660
QY 661 LKXDFSEPLNLD 672
DB 661 LESDFNEELTLD 672

RESULT 4
A44276
noncapsid protein NS1 - parvovirus LuIII
C:Species: parvovirus LuIII
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C:Accession: A44276
R:Difffoot, N.; Chen, K.C.; Bates, R.C.; Lederma, M.
Virology 192, 339-345, 1993
A:title: The complete nucleotide sequence of parvovirus LuIII and localization of a unique
A:Reference number: A44276; MUID:93297126; PMID:8517025
A:Accession: A44276
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-668 <DIF>
A:Cross-references: UNIPROT:P36311; GB:M81888
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match 91.3%; Score 3322; DB 1; Length 668;
Best Local Similarity 91.4%; Pred. No. 6.3e-233;
Matches 614; Conservative 23; Mismatches 31; Indels 4; Gaps 2;

QY 1 MAGNAYSDEVLGATNMLKEKSNQEVFSFVKENNVOLNGKDIGNSYKKELODELKSLQ 60
DB 1 MAGNAYSDEVLTGTTNMLKDKSNQEVFSFVKENNVOLNGKDIGNSYKKELODELKSLQ 60
QY 61 RGAETTDQSEDMEWETTVDMTKKQVIFDLSLVKKCLFEVLNTKNIFFPGDVNWFVQHEW 120
DB 61 RGAETTDQSEDMEWESSDELTKQVIFDLSLVKKCLFEVLNTKNIFFPGDVNWFVQHEW 120
QY 121 GKQDQGHCHVLIGKDFSOAQGWRRQLNLYWSRWLVTAACNQLTPAERIKLREIAEDN 180
DB 121 GKQDQGHCHVLIGKDFSOAQGWRRQLNLYWSRWLVTAACNQLTPAERIKLREIAEDQ 180
QY 181 EWTLLTYKHQTKQDYKCVKFCGNMAYYFLTKKKISTSPPRDGGYFLSSDSGWKTNFL 240
DB 181 EWTLLTYKHQTKQDYKCVKFCGNMAYYFLTKKKISTSPPRDGGYFLSSDSGWKTNFL 240
QY 241 KEGERHLVSKLYTDDMRPETVETTTTAQETKGRIOTKKEVSIKTYLKLVLHVRVTSPE 300
DB 241 KEGERHLVSKLYTDDMRPETVETTTTAQETKGRIOTKKEVSIKTYLKLVLHVRVTSPE 300
QY 301 DNMWMPDYSIENMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAEKSLTNFSLPD 360
DB 301 DNMWMPDYSIENMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAEKSLTNFLLAD 360
QY 361 TRACRIFAFHGMWYKVKCHAIACVNLNRQGGKRNITVLFHGPASTGKSIIAQIAQAVGNVG 420
DB 361 TRCIRIFAFHGMWYKVKCHAIACVNLNRQGGKRNITVLFHGPASTGKSIIAQIAQAVGNVG 420
QY 421 CYNAAVNFPPNDCTNKNLIWVEEAGNFGQOVNQFKAICSGQIRIDQKKGSKQIEPTP 480
DB 421 CYNAAVNFPPNDCTNKNLIWVEEAGNFGQOVNQFKAICSGQIRIDQKKGSKQIEPTP 480
```

```
DB 421 CYNAAVNFPPNDCTNKNLIWVEEAGNFGQOVNQFKAICSGQIRIDQKKGSKQIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNWPMICAWLVK 540
DB 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNWPMICAWLVK 540
QY 541 NGYSTMASYCAKWKGVDPDSENWAEKPYPTPINLLGSARSPTTTPKSTPLSQNYALTPL 600
DB 541 NGYSTMASYCAKWKGVDPDSENWAEKPYPTPINLLGSARSPTTTPKSTPLSQNYALTPL 600
QY 601 ASLEDLALEPWPSTPNTPVAGTAEQTONTGSAKQACODGQLSPTWSEIEEDLRACFGAEP 660
DB 601 ASLEDLALEPWPSTPNTPVAGTAEQTONTGSAKQACODGQLSPTWSEIEEDLRACFGAEP 660
QY 661 LKXDFSEPLNLD 672
DB 661 WKSDESLQPLNLD 668

RESULT 5
UYPVME
noncapsid protein NS1 - mink enteritis virus (strain Abaehiri)
C:Species: mink enteritis virus, MEV
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: A38350
R:Kariatsumari, T.; Horiuchi, M.; Hama, E.; Yaguchi, K.; Ishiguro, N.; Goto, H.; Shinag
J. Gen. Virol. 72, 867-875, 1991
A:title: Construction and nucleotide sequence analysis of an infectious DNA clone of the
A:Reference number: A38350; MUID:91202123; PMID:2016597
A:Accession: A38350
A:Molecule type: DNA
A:Residues: 1-668 <XAR>
A:Cross-references: UNIPROT:P27438; GB:D00765; NID:g222435; PIDN:BAA00662.1; PID:g222436
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match 72.8%; Score 2647; DB 1; Length 668;
Best Local Similarity 73.1%; Pred. No. 5.9e-184;
Matches 491; Conservative 71; Mismatches 104; Indels 6; Gaps 5;

QY 1 MAGNAYSDEVLGATNMLKEKSNQEVFSFVKENNVOLNGKDIGNSYKKELODELKSLQ 60
DB 1 MSGNQTEEVMEGVNMLKHAENEAFSFKCNVQLNGKDVHNNYTKFIQNEELTSLI 60
QY 61 RGAETTDQSDS--EDMEWETTVDMTKKQVIFDLSLVKKCLFEVLNTKNIFFPGDVNWFVQH 118
DB 61 RGAETAMDQTEEBEEMDESEVDSLAKKQVQTFDALIKKCLFEVFSKNIEPNECWEIFQH 120
QY 119 EWGKQDQGHCHVLIGKDFSOAQGWRRQLNLYWSRWLVTAACNQLTPAERIKLREIAE 178
DB 121 EWGKQDQGHCHVLHSHKNLQQAATGKWLRRQMNMYWSRWLVTLCSVNLTPTEKIKLREIAE 180
QY 179 DNEWVTLTYKHQTKQDYKCVLFGNMIAYYFLTKKKISTSPPRDGGYFLSSDSGWKTN 238
DB 181 DSEWVTLTYRHQTKQDYKCVNWFHGNMAYYFLTKKKI--VHMTKESGYFLSDSDGHWKFN 239
QY 239 FLKEGERHLVSKLYTDDMRPETVETTTTAQETKGRIOTKKEVSIKTYLKLVLHVRVTS 298
DB 240 FMKYQDRHTVSTLYTSEQMKPETVETTTTAQETKGRIOTKKEVSIKTYLRLDLVSKRVTS 299
QY 299 PEDWMMQPDYSIENMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAEKSLTNFSL 358
DB 300 PEDWMMQPDYSIENMAQPGGENLLKNTLEICTLTARTKTAFLDILEKANNKLTNFDL 359
QY 359 PDTRACRIFAFHGMWYKVKCHAIACVNLNRQGGKRNITVLFHGPASTGKSIIAQIAQAVGN 418
DB 360 ANSRTCOIFRMHGMWYKVKCHAIACVNLNRQGGKRNITVLFHGPASTGKSIIAQIAQAVGN 419
QY 419 VGCYNAANVNFPPNDCTNKNLIWVEEAGNFGQOVNQFKAICSGQIRIDQKKGSKQIEP 478
DB 420 VGCYNAANVNFPPNDCTNKNLIWVEEAGNFGQOVNQFKAICSGQIRIDQKKGSKQIEP 479
QY 479 TPVIMTNNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNWPMICAWL 538
```

Db 480 TPVIMTTNENITIVRIGCEERPEHTQPIRDRMLNKLVCPLPGDFGLVDKEWPLICAWL 539
QY 539 VKNGYSTMASYCAKWKVPDWSNWAEPKVPPTPINLGSARSPTTTPKSTPLSONYALT 598
Db 540 VKHGYSTWANYTHHWKVPEDNWAEPKIQGGI-ISPCKOLETQAANPOSQDHVLT 598
QY 599 PLASDLEDLALPEWSTPNTVAGTAETQNTGEAGSKACODGQLSPTWSEIEEDLRACFGA 658
Db 599 PLTPDVVDLALPEWSTPNTVAGTAETQNTGEAGSKACODGQLSPTWSEIEEDLRACFGA 658
QY 599 PLTPDVVDLALPEWSTPNTVAGTAETQNTGEAGSKACODGQLSPTWSEIEEDLRACFGA 658
Db 599 PLTPDVVDLALPEWSTPNTVAGTAETQNTGEAGSKACODGQLSPTWSEIEEDLRACFGA 658
QY 659 EPLKXDFSEPLN 670
Db 657 EQLEEDFRDLD 668

RESULT 6

UYPVFP
noncapsid protein NS1 - feline panleukopenia virus (strain 193)
C/Species: feline panleukopenia virus, FPLV
C/Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004
C/Accession: A36608
R/Martyn, J.C.; Davidson, B.E.; Studdert, M.J.
J. Gen. Virol. 71, 2747-2753, 1990
A/Title: Nucleotide sequence of feline panleukopenia virus: comparison with canine parvovirus
A/Reference number: A36608; MUID:91073139; PMID:2174965
A/Accession: A36608
A/Molecule type: DNA
A/Residues: 1-668 <MAR>
A/Cross-references: UNIPROT:P24842; GB:X55115; NID:G60863; PIDN:CAA38910.1; PID:G60864
C/Superfamily: parvovirus noncapsid protein
C/Keywords: noncapsid protein

Query Match 72.7%; Score 2644; DB 1; Length 668;
Best Local Similarity 73.1%; Pred. No. 9.8e-184;
Matches 491; Conservative 71; Mismatches 104; Indels 6; Gaps 5;

QY 1 MAGNAYSDEVLGATNWLKESNQVFSVFKNENVLNGKDIGNSYKKELODELKSLQ 60
Db 1 MSGNQYTEEVMEGVNWLKHAENAFSFKCDNVQNGKDVRRNNTKPIQNEELTSLI 60
QY 61 RGAETTWDS--EDMEWETTVDKTKOVFIKCLFEVLNTKNIFFGVNMFVQH 118
Db 61 RGAETTWDS--EDMEWETTVDKTKOVFIKCLFEVLNTKNIFFGVNMFVQH 118
QY 119 EWGKDCGWHCHVLGGKDFQAGKWRRLQNVNWSRLVATCNVLTTPAERIKLREIAE 178
Db 121 EWGKDCGWHCHVLGGKDFQAGKWRRLQNVNWSRLVATCNVLTTPAERIKLREIAE 180
QY 179 DNEWVTLTYKHQTKDYKCVLFGNMIAFYFLTKKISTSPRGGYFLSDSGWKFN 238
Db 181 DSEWVTLTYRHQTKDYKCVLFGNMIAFYFLTKKISTSPRGGYFLSDSGWKFN 239
QY 239 FLKEGERHLVSKLYTDDMRPETVTTTAQETKRGRIOTKKEVSIKTLKELVHKRVTS 298
Db 240 FMKYQDRHTVSTLYTEQMPETVTTTAQETKRGRIOTKKEVSIKTLKELVHKRVTS 299
QY 299 PEDMMWQPSYIEMMAQPGENLLKNTLEICTLTARTKTAFLILEKETSCLTNFSL 358
Db 300 PEDMMWQPSYIEMMAQPGENLLKNTLEICTLTARTKTAFLILEKADNTKLTNFDL 359
QY 359 PDTRACRIFAFHGNVYKVCCHAIICCVLNROGKRNVTFLFHPASTGKSIQAIAQAVGN 418
Db 360 ANSRCTQIFRMHGNWLVKCHAIACVNLNROGKRNVTFLFHPASTGKSIQAIAQAVGN 419
QY 419 VGCYNAANVPNDCTNKNLIWVEAGNFGQVNOFKAI CSQTIRIDQKGGSKQIEP 478
Db 420 VGCYNAANVPNDCTNKNLIWVEAGNFGQVNOFKAI CSQTIRIDQKGGSKQIEP 479
QY 479 TPVIMTTNENITIVRIGCEERPEHTQPIRDRMLNKLVCPLPGDFGLVDKEWPLICAWL 538
Db 480 TPVIMTTNENITIVRIGCEERPEHTQPIRDRMLNKLVCPLPGDFGLVDKEWPLICAWL 539

QY 539 VKNGYSTMASYCAKWKVPDWSNWAEPKVPPTPINLGSARSPTTTPKSTPLSONYALT 598
Db 540 VKHGYSTWANYTHHWKVPEDNWAEPKIQGGI-ISPCKOLETQAANPOSQDHVLT 598
QY 599 PLASDLEDLALPEWSTPNTVAGTAETQNTGEAGSKACODGQLSPTWSEIEEDLRACFGA 658
Db 599 PLTPDVVDLALPEWSTPNTVAGTAETQNTGEAGSKACODGQLSPTWSEIEEDLRACFGA 658
QY 659 EPLKXDFSEPLN 670
Db 657 EQLEEDFRDLD 668

RESULT 7

UYPVCP
noncapsid protein NS1 - canine parvovirus (strain N)
C/Species: canine parvovirus, CPV
C/Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
C/Accession: A29962
R/Reed, A.P.; Jones, E.V.; Miller, T.J.
J. Virol. 62, 265-276, 1988
A/Title: Nucleotide sequence and genome organization of canine parvovirus.
A/Reference number: A29962; MUID:88062992; PMID:2824850
A/Accession: A29962
A/Molecule type: DNA
A/Residues: 1-668 <REE>
A/Cross-references: UNIPROT:P12929; EMBL:M19296; NID:G333438; PIDN:AAA67459.1; PID:G33334
C/Superfamily: parvovirus noncapsid protein
C/Keywords: noncapsid protein

Query Match 72.6%; Score 2642; DB 1; Length 668;
Best Local Similarity 73.1%; Pred. No. 1.4e-183;
Matches 491; Conservative 71; Mismatches 104; Indels 6; Gaps 5;

QY 1 MAGNAYSDEVLGATNWLKESNQVFSVFKNENVLNGKDIGNSYKKELODELKSLQ 60
Db 1 MSGNQYTEEVMEGVNWLKHAENAFSFKCDNVQNGKDVRRNNTKPIQNEELTSLI 60
QY 61 RGAETTWDS--EDMEWETTVDKTKOVFIKCLFEVLNTKNIFFGVNMFVQH 118
Db 61 RGAETTWDS--EDMEWETTVDKTKOVFIKCLFEVLNTKNIFFGVNMFVQH 120
QY 119 EWGKDCGWHCHVLGGKDFQAGKWRRLQNVNWSRLVATCNVLTTPAERIKLREIAE 178
Db 121 EWGKDCGWHCHVLGGKDFQAGKWRRLQNVNWSRLVATCNVLTTPAERIKLREIAE 180
QY 179 DNEWVTLTYKHQTKDYKCVLFGNMIAFYFLTKKISTSPRGGYFLSDSGWKFN 238
Db 181 DSEWVTLTYRHQTKDYKCVLFGNMIAFYFLTKKISTSPRGGYFLSDSGWKFN 239
QY 239 FLKEGERHLVSKLYTDDMRPETVTTTAQETKRGRIOTKKEVSIKTLKELVHKRVTS 298
Db 240 FMKYQDRHTVSTLYTEQMPETVTTTAQETKRGRIOTKKEVSIKTLKELVHKRVTS 299
QY 299 PEDMMWQPSYIEMMAQPGENLLKNTLEICTLTARTKTAFLILEKETSCLTNFSL 358
Db 300 PEDMMWQPSYIEMMAQPGENLLKNTLEICTLTARTKTAFLILEKADNTKLTNFDL 359
QY 359 PDTRACRIFAFHGNVYKVCCHAIICCVLNROGKRNVTFLFHPASTGKSIQAIAQAVGN 418
Db 360 ANSRCTQIFRMHGNWLVKCHAIACVNLNROGKRNVTFLFHPASTGKSIQAIAQAVGN 419
QY 419 VGCYNAANVPNDCTNKNLIWVEAGNFGQVNOFKAI CSQTIRIDQKGGSKQIEP 478
Db 420 VGCYNAANVPNDCTNKNLIWVEAGNFGQVNOFKAI CSQTIRIDQKGGSKQIEP 479
QY 479 TPVIMTTNENITIVRIGCEERPEHTQPIRDRMLNKLVCPLPGDFGLVDKEWPLICAWL 538
Db 480 TPVIMTTNENITIVRIGCEERPEHTQPIRDRMLNKLVCPLPGDFGLVDKEWPLICAWL 539
QY 539 VKNGYSTMASYCAKWKVPDWSNWAEPKVPPTPINLGSARSPTTTPKSTPLSONYALT 598
Db 540 VKHGYSTWANYTHHWKVPEDNWAEPKIQGGI-ISPCKOLETQAANPOSQDHVLT 598

Db 421 YNAANVNFNDCTNKNLWIEEAGNFQVNOQFRAICSGQTIRIDQKGGSKQIEPTPV 480
Qy 482 IMTNTNITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNWPMICAWLVKN 541
Db 481 IMTNTNITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNWPMICAWLVKN 540
Qy 542 GYQSTMASYCAKMGVDPWSEWAEKVPPTPINLIGSARSFPTPKSTPLSQNTALPLA 601
Db 541 GYQATMASYMHGWNVDWSEKWEKQWQTPINPTDSQIS-TSVKTSPADNNVAAPFIQ 599
Qy 602 SDLE-DLALPWSPTNTVAGTAETQNTGEAGSKACQDQQL---SPTWSEIEEDLRACFG 657
Db 600 EDLDLALPWSPTNTVAGTAETQNTGEAGSKACQDQQL---SPTWSEIEEDLRACFG 652

RESULT 10

UVPVIF
noncapsid protein NS1 - feline panleukopenia virus (fragment)
C:Species: feline panleukopenia virus, FPLV
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 09-Jul-2004
C:Accession: A03697
R:Carlson, J.; Rushlow, K.; Maxwell, I.; Maxwell, F.; Winston, S.; Hahn, W.
J. Virol. 55, 574-587, 1985
A:Title: Cloning and sequence of DNA encoding structural proteins of the autonomous parvovirus
A:Reference number: A03697; UID:85265017; PMID:2991581
A:Accession: A03697
A:Molecule type: DNA
A:Residues: 1-392 <CAR>
A:Cross-references: UNIPROT:P06431; EMBL:M10824; NID:g333474; PIDN:AAA47160.1; PID:g333474
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match 44.6%; Score 1621.5; DB 1; Length 392;
Best Local Similarity 77.5%; Pred. No. 6.7e-110; Mismatches 53; Indels 3; Gaps 3;
Matches 306; Conservative 33

Qy 276 IQTKVESIKTKLKVHRTSPEDMMQPDSEYIEMMAQPGGKNTLEICTITLA 335
Db 1 IQTKVESIKTKLKVHRTSPEDMMQPDSEYIEMMAQPGGKNTLEICTITLA 60
Qy 336 RTKTAFLILEKATSKLTNPSLDPTRACIPAFHGWYVVKVCHACVNLNROGKNTV 395
Db 61 RTKTAFLILEKANTKLTNFDLANSRTCQIFRMHGNWIKVCHACVNLNROGKNTV 120
Qy 396 LFHGPASTGKSIIAQAIAQAVGNCYNAANVPFNDCTNKNLIWVEEAGNFQOQVNOF 455
Db 121 LFHGPASTGKSIIAQAIAQAVGNCYNAANVPFNDCTNKNLIWVEEAGNFQOQVNOF 180
Qy 456 KAICSGQTIRIDQKGGSKQIEPTPVIMTNNITVVRIGCEERPEHTQPIRDRMLNHL 515
Db 181 KAICSGQTIRIDQKGGSKQIEPTPVIMTNNITVVRIGCEERPEHTQPIRDRMLNHL 240
Qy 516 THLHPGDFGLVDKNWPMICAWLVKGYOSTMASYCAKMGKVPDWSNWAEPKVPPTINL 575
Db 241 VCKLPGDFGLVDKNWPMICAWLVKGYOSTMASYCAKMGKVPDWSNWAEPKVPPTINL 300
Qy 576 LGSARSFPTPKSTPLSQNTALPLADLEALPWSPTNTVAGTAETQNTGEAGSKA 635
Db 301 PG-CKDLETOASNPQSDHVLPTLTPVDVLDALPWSPTNTVAGTAETQNTGEAGSKA 357
Qy 636 QDQQLSPTWSEIEEDLRACFGAEPLKQFSEPLN 670
Db 358 HKDVQASPTWSEIEADLRAIFTSQLEEDFRDLD 392

RESULT 11

S41439
gene NS-1 protein - Aleutian mink disease virus
C:Species: Aleutian mink disease virus
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S41439
R:Gottschalk, E.; Alexandersen, S.; Storgaard, T.; Bloom, M.E.; Aasted, B.
submitted to the EMBL Data Library, January 1994

submitted to the EMBL Data Library, January 1994
A:Description: Sequence comparison of the non-structural genes of four different types of
A:Reference number: S41434
A:Accession: S41439
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-641 <GOT>
A:Cross-references: UNIPROT:Q65017; EMBL:Z29576
C:Genetics:
A:Introns: 586/1
C:Superfamily: parvovirus noncapsid protein

Query Match 15.1%; Score 550; DB 2; Length 641;
Best Local Similarity 25.8%; Pred. No. 7.8e-32;
Matches 178; Conservative 108; Mismatches 262; Indels 142; Gaps 25;

Qy 18 KEKSNQVFSFVFNENVQLNGKDIGNSYKKELOBDE-----LKSQORGAETWD 68
Db 8 EQRRLQDLVYQLKKEIN--DGEQVAVLFOQKTYTDKDKNPKTRATPLRTTSSDLRLAFD 64
Qy 69 QSEDMEWETTVDSEWK-----KQVFIQSLVKKCLFEVLNATKNIFFPGDVNWFVQ 117
Db 65 SIEE-NUTASNEQNTNGINFCKLTGLKTLILLDDKHVSHRWD--NNK-----VNLIWQ 115
Qy 118 HEMGKQDQGWCHVLIG-----GKDFSOAQGWRRRLN-----VWMSRLVTACNVQ- 164
Db 116 IEKGKTKQFHHCHCLGVDFDRNEDPKDVQKSLG-WFMKRLNKDLALIYNSH---HCDIQD 170
Qy 165 -LTPAEIKURETAENENWTLTYKHQKQDYTKCVLFGNMIAYYFLTKKISTSPPR 223
Db 171 IKDPEDRAKMLKVMVEDGPTKPKYFNKQTKDYKPNVHLRDYTFYLFNKDKINTDSM- 229
Qy 224 DGGYFLLSDSGWKTNFLKGERHLVSKLYTDDMRPTEVTV-----TTAQ 269
Db 230 -DGYFAGNGIIVDN-LTKKERTLRMYLDEOSSDIMANDIDWEDQDQAPKVTDQDSA 287
Qy 270 ETKRG-----RIOTKKEV-----SIKTKLKVHRTSPEDMMQPD 307
Db 288 TTKTGTSLIWKSCATKVTSKVANPVQPSKLYSAQNTLDALFVNGCFTPEDMIKQS 347
Qy 308 DSVIEMMAQPGGKNTLEICTITLARTKTAFLILE--KAETSKLTNFSLPDTRACR 365
Db 348 DKYLELSLEPSPGQKINTLLHMQVKTSTMTAFDCIIRFNEEDDKPLATIKDM---- 403
Qy 366 IFAPHGWNVVKVCHACVNLNROGKNTVLFHGPASTGKSIIAQAIAQAVGNCVNA 425
Db 404 -----GLNEQYLKVVCTILTKQGGKRGCIWFGPGGTGKTLASLICKATVNYGMVTS 458
Qy 426 NVNFPNDCTNKNLIWVEEAGNFQOQVNOFKAICSGQTIRIDQKGGSKQIEPTPVIMT 485
Db 459 NPNFPMTDCGNRIIWAEECGNLGNWVEDFKAITGGGVKVDTPKNQPSIKGC-VIVTS 517
Qy 486 NENITVVRIGCEERPEHTQPIRDRMLNHL-----THHLPGDFGLVDKNWPMICAWL 538
Db 518 NTNITKTVGCVETNAHAEPKQRMKIKRCMTKINPKTKITFG-----MLKRWL 566
Qy 539 VKNGYQSTMASYCAKMGKVP-DWSNWAEPKVPPTINLIGSARSFPTPKSTPLSQNTAL 597
Db 567 -----NTWDRQPIQLSHPEPELYLET-----TGNSSATATKNTGNSQ---- 605
Qy 598 TPLASDLALPWSPTNTVAGTAETQNT 627
Db 606 PTTAKSAESVNTENCDDTPKRGASSVPPKQH 635

RESULT 12

S41434
gene NS-1 protein - Aleutian mink disease virus
C:Species: Aleutian mink disease virus
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S41434
R:Gottschalk, E.; Alexandersen, S.; Storgaard, T.; Bloom, M.E.; Aasted, B.
submitted to the EMBL Data Library, January 1994

A;Description: Sequence comparison of the non-structural genes of four different types of parvovirus
A;Reference number: S41434
A;Accession: S41434
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-641 <GOT>
A;Cross-references: UNIPROT:Q65020; EMBL:Z29577
C;Genetics:
A;Introns: 586/1
C;Superfamily: parvovirus noncapsid protein

Query Match 15.08; Score 544.5; DB 2; Length 641;
Best Local Similarity 25.28; Pred. No. 2e-31;
Matches 182; Conservative 110; Mismatches 260; Indels 171; Gaps 26;

Qy 18 KEKSNQVVFVFKQENNVQLNGKDIGNSVKKELQEDF-----LKSILQGAETWD 68
Db 8 EQRLQLBFE-KFYTE--VADGGLAWLFOQKTYTDKDNKPTKATPLRTTSSDLRLAFD 64
Qy 69 QSEDMEWETTVDKTKQVIFDLSLVKKCLFEVLNTKNI-----FPGDVNVFQHEWKGKD 123
Db 65 SIEE-TLKTSNQCLTNNDINFCCLTLGKTL--VLLDKHVRSHRDANKVNFVQVEKGKT 121
Qy 124 QGWHCHVLIG-----GKDFSOAQGWRRQLN-----VYWSRWLVTCNVOLTPAERIK 172
Db 122 QQPHIHCCLGYPDKDEDSKDVQKSLG-WFKKLNKDLAVIWSNHHCDIQIGQSDGRADN 180
Qy 173 LREIAEDNEWVTLTYKHGKOTKDYTKCVLFGNMIAYFLTKKISTSPRDG--GYFLS 230
Db 181 LKTIWIEGPT-KPKYKFNKQTKQDYNKPTNLRDYLILYFNKDKIT-----QEGMDGYAS 235
Qy 231 SDSGKWNFLKEGERHLVSKLYTDDM-----RPETVETTVAQETKRG 274
Db 236 GNGGIIDN-LTNKERKALRWYLDQSQGILDEDIDWEDSQSAPKVTQDTSATSKTGS 294
Qy 275 -----RQTKKEVSI-----KTKLKLHVRKVTSPEDMMQPDSEYEMM 314
Db 295 LVMSKCATKVTSKVEVAIPVKQPSKQWTSQAQNTLDDLYMFGCFTPEDMILKMSDRYLEMS 354
Qy 315 AQPGENLLKNTLEICTLTARTKTAFDLILE--KAETSKLTNPSLPDTRACRIFAFHGW 372
Db 355 LEPNGAKINTLLHMNQVTSNLTAFECIKFNEEEDKPLLDITIKDM-----GL 405
Qy 373 NYVKVCHACVCLNROGKRNVTFLHFGPASTGKSIIAQAIAQAVGNVGCYNAANVPFPN 432
Db 406 NEQHLKKVLCTILTQSGKRGCVFYGPGGTGKTLASLICKAVVNGVMTTSNPNFPWT 465
Qy 433 DCTNNLIWVEEAGNFQQVNOFKAIKSGQTIRIDQKSGSKQIEPTPVIMTNENITVV 492
Db 466 DCGNRNIWAEECGNIGNYVEDFKAITGGDVKVDTKNKPQSIKGS-VIVTSNTNITKV 524
Qy 493 RIGCEERPEHTOPTIRDRMLNIHL-----THLPGDFGLVDKNEWPMICAWLVKNGYQS 545
Db 525 TVGCVETNVIAEPLQKRMKIRCMKPNPKVTPTG-----MLKWTI----- 566
Qy 546 TMASYCAKWKGVDPWSENWAEKVPPTPINLIGSARSPTTPKSTPLSQNYALTPLASDLE 605
Db 567 -----STWDRIIP-----IKLSHEMP 581
Qy 606 DLALPSTNTPVAGTAETQNTGEAGSKACQDQLSPTWSEIIEEDLRACFAGPLKDF 665
Db 582 ELYLET-SGPNSS--SATTATKST-----GNLQPTTAETAESVNTANCDDTPKRGAS 628
Qy 666 SEP 668
Db 629 SVP 631

RESULT 13
UYVPVAP
noncapsid protein NS1 - Aleutian mink disease virus (strain ADV-G)
N;Alternate names: left-ORF protein
C;Species: Aleutian mink disease virus

C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C;Accession: A36760; A35529
R;Bloom, M.E.; Alexandersen, S.; Perryman, S.; Lechner, D.; Wolfbarger, J.B.
J. Virol. 62, 2903-2915, 1988
A;Title: Nucleotide sequence and genomic organization of Aleutian mink disease parvovirus
A;Reference number: A36760; MUID:88275062; PMID:2839709
A;Accession: A36760
A;Molecule type: DNA
A;Residues: 1-620 <BL2>
A;Cross-references: UNIPROT:P24030; EMBL:M20036
C;Superfamily: parvovirus noncapsid protein
C;Keywords: noncapsid protein

Query Match 14.98; Score 543; DB 1; Length 620;
Best Local Similarity 26.68; Pred. No. 2.4e-31;
Matches 157; Conservative 104; Mismatches 229; Indels 100; Gaps 18;

Qy 18 KEKSNQVVFVFKQENNVQLNGKDIGNSVKKELQEDLKSILQGAETWDQSB-DMEWE 76
Db 38 EQRLQLBLYVOLKEIN---DGGVAVLFOQKTYTDKDNKPTKATPLRTTSSDLRLAFD 94
Qy 77 TTVDEMTKKQVIFDLSLVKKCLFEVLNTKNI-----PGDVNVFQHEWKGQDGV 126
Db 95 SIEENLTASNEHLTNNEINFCKLTGLTKTLLIDKHVKSHERWDSNKVNLINQIEKGTQOF 154
Qy 127 HCHVLIG-----GKDFSOAQGWRRQLN-----VYWSRWLVTCNVQ--LTPAERIKL 173
Db 155 HIHCLGLGYFDKNEPDKDVQKSLG-WFKRLNKLAVIYSN---HHCDIQIDKPEDRAKN 210
Qy 174 RETAEDNEWVTLTYKHGKOTKDYTKCVLFGNMIAYFLTKKISTSPRDGQVFLSSDS 233
Db 211 LKVMIEDGPTKPKYKFNKQTKQDYNKPEVHLRDYTFIYFNKDKINTDSM--DGYPAAAGNG 268
Qy 234 GWKTNFLKEGERHLVSKLYTDDMRPETVETV-----TTAETKRG----- 274
Db 269 GIVDN-LTNKERKTLRWYLDQSQSDIMDANIDWEDQDAPKVTQDTSATTKGTSLIW 327
Qy 275 -----RQTKKEV-----SIKTKLKLHVRKVTSPEDMMQPDSEYEMMAQP 317
Db 328 KSCATKVTSKVEVANPVQPSKELYSQAQSTLDALFNVCFTPEDMIILKQSKYLELSLEP 387
Qy 318 GGENLLKNTLEICTLTARTKTAFDLILE--KAETSKLTNPSLPDTRACRIFAFHGWYV 375
Db 388 NGPQKINTLLHMNQVTSNLTAFDCLIKFNEEEDKPLLATIKDM-----GLNEQ 438
Qy 376 KVCHACVCLNROGKRNVTFLHFGPASTGKSIIAQAIAQAVGNVGCYNAANVPFPNDCT 435
Db 439 YLKKVLCITLTQSGKRGCVFYGPGGTGKTLASLICKATVNGVMTTSNPNFPWTDG 498
Qy 436 NKNLWVEEAGNFQQVNOFKAIKSGQTIRIDQKSGSKQIEPTPVIMTNENITVVIR 495
Db 499 NRNIWAEECGNFGNVVEDFKAITGGDVKVDTKNKPQSIKGC-VIVTSNTNITKVTVG 557
Qy 496 CEEPEHTOPTIRDRMLNIHL-----THLPGDFGLVDKNEWPMICAWL 538
Db 558 CVETNAHAEPKQRMKIRCMKTNPKTKITPG-----MLKRWL 596

RESULT 14
S41861
gene NS-1 protein - Aleutian mink disease virus
C;Species: Aleutian mink disease virus
C;Date: 15-Sep-1994 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
C;Accession: S41861; S41436
R;Gottschalk, E.; Alexandersen, S.; Storgaard, T.; Bloom, M.E.; Aasted, B.
submitted to the EMBL Data Library, January 1994
A;Description: Sequence comparison of the non-structural genes of four different types of parvovirus
A;Reference number: S41434
A;Accession: S41861
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-641 <GOT>
A;Cross-references: UNIPROT:Q65023; EMBL:X77085; NID:9452597; PIDN:CAA54355.1; PID:945252

C;Genetics:

A;Introns: 586/1

C;Superfamily: parvovirus noncapsid protein

Query Match 14.7%; Score 533.5; DB 2; Length 641;
Best Local Similarity 25.7%; Pred. No. 1.2e-30;
Matches 178; Conservative 106; Mismatches 250; Indels 159; Gaps 27;

QY 18 KEKSNQEVFSFVKENNVQLANGKIDGWSYKELQEDS-----LKSILQRGAETIWD 68
DB 8 EQLRLDLYTQLKE---VADGGLAWLFQOKTYTDKDNKTKATPPURITSSDLRLAFD 64
QY 69 QSEDM-----EWETTVDMT-----KKQVFIPLSLVKKCLFEVLNTKNIFPGDVNMFVQ 117
DB 65 SIEETLIASNEW-LTKDBINFCKLTGLKTLVLVDKHKVSHRWA-----DKINFIQ 115
QY 118 HEMGKQGWCHVLIG-----GKFSQAQGWRRQLNVYWRWLVTACNVOLTPAER 170
DB 116 IEKGTQHFHTHCLGYFDKNEPKDVQKSLG-WLIKLN---RDAAIFSNHHCQDQD 170
QY 171 IKLEIAEDN--EWV-----TLLTYKHKTQKDYTKCVLFGNMIAYFLTKKISTSPPR 223
DB 171 IKDPEAKANNLVIEDGTPYKYKQTKQDTNKPVTLSDYVILYFNKDKIH-----K 226
QY 224 DG--GYFLSSDSGWKTNPLKEGERHLVSKLYTDDMRPETVE-----TTVTT 267
DB 227 EGMGYAAGNGGLIDN-LTNKERKALKRMYLDEQSSDMDADIDWEDGQDAPKVTQD 285
QY 268 AQETKRG-----RIQTKKV-----SIKTLKELVHKRVTSPEWMM 305
DB 286 SATSKTGLSLWKSCATKVTSKKEVANPVOQPSKKLYSAQNTLDALFNVGCTPEDMIK 345
QY 306 QPDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFLDLE---KAETSCLTNFSLPDTRA 363
DB 346 QSDXYLESLSPNGPKINTLLHNNQVKTSTMTAFDCIKFNEEDDKPLLATIKM-- 403
QY 364 CRIFAFHGNVYVCHACCVLNRGGKRNVLPHGPASTGKSIIAQIAQAVNGVCYN 423
DB 404 -----GLNEQYLKVLCTLTQGGKRGCIWFYGGTGKTLASLICKATVNYGMVT 456
QY 424 AANVNFENDCTNKLIVWEAGNFGQVNOFKAI CSQTIRIDQKSGSKQIEPTVIM 483
DB 457 TSNPNFPMWDCGNRIIWAEECGNLGNVEDFKAITGGGVKVDTKNKPQSIKGC-VIV 515
QY 484 TTENITVTRIGCERPEHTOPIRDRMLNIHL-----THLPGDGLVDKNEWPMICA 536
DB 516 TSNITITVTVGCVETNAHAEPKQRMKICMKTINPKTKITPG-----MLKR 564
QY 537 WLVRNGYOSTWASYCAKWKV-----DWSENWAEKVPVTPINLLGSARGPP-TTPKSTP 590
DB 565 WL-----NTWDRQPIQLSHEMPELYLETTGPN-----SSATSATKTTGNSQP 606
QY 591 LSQNYALTPLASDLELALPWPSTNTPVAGTA 623
DB 607 TTAATSVSTADCD-----TPKRGAS 628

RESULT 15

S52209

noncapsid protein NS1 - muscovy duck parvovirus

N;Alternate names: nonstructural protein NS1; NS1 protein

C;Species: muscovy duck parvovirus

C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004

C;Accession: S52209

R;Zadori, Z.; Erdei, J.; Nagy, J.; Kisari, J.

submitted to the EMBL Data Library, September 1993

A;Reference number: S52209

A;Accession: S52209

A;Molecule type: DNA

A;Residues: 1-626 <ZAD>

A;Cross-references: UNIPROT:Q81288; EMBL:X75093; NID:g609091; PIDN:CAA52983.1; PID:g60909

A;Experimental source: strain FM

C;Genetics:

A;Gene: NS1

C;Superfamily: parvovirus noncapsid protein

C;Keywords: noncapsid protein

Query Match 12.5%; Score 455.5; DB 2; Length 626;
Best Local Similarity 25.2%; Pred. No. 5.4e-25;
Matches 164; Conservative 82; Mismatches 264; Indels 141; Gaps 22;

QY 75 WEET-----VDEMTKKQYFIPLSLVKKCLFEVLNTKNIF-----PCDVNWFVQHEWKG 122
DB 45 WEPFGIWNHEVNLFWMTLADKI-----KNIFIQRMNQFNODETFFQLSEGS 93
QY 123 DQWHCHVLJGGKDFSOAQGWRRQLNVYWRWLVTACNVQLTPAERIK---LRETAED 179
DB 94 EY-IHLHAVCPGECRSFVLGRY-----MSQIKDSILRDVVEG 129
QY 180 N-----EWVTLITVYKHQTKDYTKCVLFGNMIAYFLTKKISTSPRGGVFLSSDSG 234
DB 130 KQVKIPDWFISITTKRGQNKVTAA-----YILHVLIPKKQ-----PELQWAFNNPLF 179
QY 235 WKTNFKEGEGERHLVSKLYTDDMRPETVETVTTTAQETKRGRIQTKKEVSIKTLKELVHK 294
DB 180 TAAALCLQKQELDLAPQESSEMNNAVQEDQAAPL-----ISNRAAKNYSNLVDWLIEM 235
QY 295 RVTSPEWMMQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFLDLEKAETSCLT 354
DB 236 GITSEKQWLTENKESYRSFOATSSNNRQVKALENAPAEMLLTATDYLIGKDPVLDT 295
QY 355 NLSLPDTRACRIFAFHGNVYVCHACCVLNRGGKRNVLPHGPASTGKSIIAQIAQ 414
DB 296 K-----NRIYQILKANNYPQYVGVSLCGWVKREFNKRNAIWLPGPATTKTNTIAEIAH 350
QY 415 AVGNVGCYNAANVFPENDCTNKLIVWEAGNFGQVNOFKAI CSQTIRIDQKSGSK 474
DB 351 AVPFYGCNVNTNENFPENDCDVKMLIWEEGKWNKVVESAKAILLGSASVRVDQCKGVS 410
QY 475 QIEPTVIMTTNENITVTRIGCERPEHTOPIRDRMLNIHLTHLPGDGLVDKNE---- 530
DB 411 CIEPTVITNTDMCMIVDGNSTTMEHRIPLERMFQIVLSHLEGNFGKISKKEVKEF 470
QY 531 --W-----PMICAWLVKNGYQSTMASYCAKWKVDPWSEN-----WAEPK----- 568
DB 471 FKWANDNLVPVSEFVKPTNEQTKLTE-----PVPERANEPSEPPKIWAPPTREELEE 524
QY 569 -----VPTPINLLGSARSPTTPKSTPLSQNYALTPLASDLELALPWPSTNPT 617
DB 525 LRASPELFSVAFLP-----SSDTPSPKRTKRGYQVRCAMHSL-DNSMNVFECLEC 576
QY 618 PVAGTAETQNTGBAGSKAC-QDGLSPTWSEIEBEDLRACFGARPLKKDFSE 667
DB 577 ERANFPFQSLGE--NFCNQHG-----WYD-----CAFCNELKDDWNE 612

Search completed: January 22, 2005, 03:42:01

Job time : 26.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:19:28 ; Search time 106.333 Seconds
(without alignments)
3636.223 Million cell updates/sec

Title: US-10-069-056-10
Perfect score: 3637
Sequence: 1 MAGNAYSDEVLGATNWLKEK.....RACFGAEPLKKDFSEPLNLD 672

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_02.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3623	99.6	672	1 VNC5 MUMIV	P0134 murine minu
2	3623	99.6	721	2 Q84365	Q84365 murine minu
3	3554	97.7	721	1 VNC5 MUMIM	P07300 murine minu
4	3554	97.7	721	2 Q84363	Q84363 murine minu
5	3552	97.7	672	2 Q83429	Q83429 mouse parvo
6	3345	92.0	672	2 Q8JVI4	Q8JVI4 rat minute
7	3336	91.7	672	2 Q8JVT8	Q8JVT8 kilham rat
8	3335	91.7	672	2 P88899	P88899 kilham rat
9	3327	91.5	672	2 Q8JVI8	Q8JVI8 rat minute
10	3326	91.4	672	1 VNC5 PAVHH	P01133 hamster par
11	3323	91.4	672	2 Q8JVT6	Q8JVT6 rat minute
12	3322	91.3	668	1 VNC5 PAVL3	P36311 parvovirus
13	3303	90.8	665	2 O71159	O71159 kilham rat
14	2935.5	80.7	671	2 O71157	O71157 rat parvovi
15	2655	73.0	668	2 P89513	P89513 feline panl
16	2654	73.0	668	2 P89516	P89516 feline panl
17	2653	72.9	668	2 P89512	P89512 feline panl
18	2653	72.9	668	2 P89515	P89515 feline panl
19	2653	72.9	668	2 P90449	P90449 feline panl
20	2653	72.9	668	2 P90484	P90484 feline panl
21	2652	72.9	668	2 P89514	P89514 feline panl
22	2652	72.9	668	2 P90472	P90472 feline panl
23	2647	72.8	668	1 VNC5 MEVA	P27438 mink enteri
24	2644	72.7	668	1 VNC5_FPV19	P24842 feline panl
25	2642	72.6	668	1 VNC5_PAVCN	P12929 canine parv
26	2637	72.5	668	2 Q84393	Q84393 canine parv
27	2614	71.9	668	2 Q70M74	Q70M74 canine parv
28	2614	71.9	668	2 CAE47433	CAE47433 canine pa
29	2485	68.3	660	1 VNC5_PAVPN	P18547 porcine par
30	2476	68.1	662	1 VNC5_PAVPK	P52502 porcine par
31	2472	68.0	662	2 Q6RED5	Q6RED5 porcine par

32	2472	68.0	662	2 Q6TPD8	Q6TPD8 porcine par
33	2472	68.0	662	2 AAQ90279	AAQ90279 porcine p
34	2472	68.0	662	2 AAR91039	AAR91039 porcine p
35	2463	67.7	662	2 Q6PS60	Q6PS60 porcine par
36	2463	67.7	662	2 AAS93262	AAS93262 porcine p
37	1933	53.1	397	2 Q993M6	Q993M6 autonomic
38	1821.5	44.6	392	1 VNC5_FPV	P06431 feline panl
39	550	15.1	641	2 Q65017	Q65017 aleutian mi
40	544.5	15.0	641	2 Q65020	Q65020 aleutian mi
41	543	14.9	590	1 VNC5 ADVG	P24030 aleutian mi
42	542	14.9	620	2 Q96607	Q96607 aleutian mi
43	533.5	14.7	641	2 Q65023	Q65023 aleutian mi
44	465.5	12.8	610	2 Q6JL80	Q6JL80 bovine aden
45	465.5	12.8	610	2 AAR26464	AAR26464 bovine ad

ALIGNMENTS

RESULT 1
VNC5 MUMIV
ID VNC5 MUMIV STANDARD; PRT; 672 AA.
AC P0134;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 03-JUL-2004 (Rel. 44, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVPL).
GN Name=NS1;
OS Murine minute virus (Murine parvovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10794;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83143341; PubMed=6298737;
RA Astell C.R.; Thomson M.; Merchlinsky M.; Ward D.C.;
RT "The complete DNA sequence of minute virus of mice, an autonomous
parvovirus.";
RL Nucleic Acids Res. 11:999-1018(1983).
RN [2]
RP HOMOLOGIZATION, AND INTERACTION WITH SYNCRIP.
RX MEDLINE=99102562; PubMed=9847309;
RA Harris C.E.; Boden R.A.; Astell C.R.;
RT "A novel heterogeneous nuclear ribonucleoprotein-like protein
interacts with NS1 of the minute virus of mice.";
RL J. Virol. 73:72-80(1999).
CC -I- FUNCTION: Seems necessary for viral DNA replication.
CC -I- SUBUNIT: Forms homooligomers. Interacts with human SYNCRIP.
CC -I- DOMAIN: The N-terminus (residues 1-275) possess a negative effect
on transactivation.
CC -I- DOMAIN: The C-terminus (residues 543-672) possess an activation
domain.
CC -I- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; J02275; AAA67109.1; -;
CC EMBL; V01115; CAA24309.1; ALT_INIT.
CC PIR; A03696; UVPVIM.
CC TRANSFAC; T02375; -;
CC InterPro; IPR001257; Parvo NS1.
CC Pfam; PF01057; Parvo NS1; I.
CC ATP-binding; DNA replication; Noncapsid protein;
CC Nonstructural protein.
CC DOMAIN 1 276 Interacts with SYNCRIP.
CC NP_BIND 399 406 ATP (Potential)
CC SEQUENCE 672 AA; 76248 MW; 50298F27662E3C1D CRC64;

```
Query Match      99.6%; Score 3623; DB 1; Length 672;
Best Local Similarity 99.7%; Pred. No. 6.6e-260;
Matches 670; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAGNAYSDEVLGATNMLKEKSNQEVFSFVKENNVQNGKDIGNWSYKKELODELKSLQ 60
Db 1 MAGNAYSDEVLGATNMLKEKSNQEVFSFVKENNVQNGKDIGNWSYKKELODELKSLQ 60
Qy 61 RGAETTWQSEDMEWETTVDEMCKQVFI FDSL VKKCLFEVLNTKNIFPGDVNMFVQHEW 120
Db 61 RGAETTWQSEDMEWETTVDEMCKQVFI FDSL VKKCLFEVLNTKNIFPGDVNMFVQHEW 120
Qy 121 GKQGWCHVLIIGKGFDSQAQGWRRQLNVLVYSLVLTACNVQLTPAERIKLREIAEDN 180
Db 121 GKQGWCHVLIIGKGFDSQAQGWRRQLNVLVYSLVLTACNVQLTPAERIKLREIAEDN 180
Qy 181 EWTLLTYKHQKQTKDYTCVLFQGNMAYFLTKKKISTSPRRGGYFLSSDSGWKTNFL 240
Db 181 EWTLLTYKHQKQTKDYTCVLFQGNMAYFLTKKKISTSPRRGGYFLSSDSGWKTNFL 240
Qy 241 KEGERHLVSKLYTDDMRPEVTETVTTAQTETKGRGRIOTKKEVSIKTTLKELVHKRVTSPE 300
Db 241 KEGERHLVSKLYTDDMRPEVTETVTTAQTETKGRGRIOTKKEVSIKTTLKELVHKRVTSPE 300
Qy 301 DMMWMPDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSPLD 360
Db 301 DMMWMPDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSPLD 360
Qy 361 TRACRIFAFHGNVYKCHAI CCVLNRQGGKRNVTLPFGPASTGKSI IAAIAQAVGNVG 420
Db 361 TRCIRIFAFHGNVYKCHAI CCVLNRQGGKRNVTLPFGPASTGKSI IAAIAQAVGNVG 420
Qy 421 CYNAAVNVFPNDCTNKNLIWVEAGNFQGVNQFKAICSGQTIIRIDQKKGSKQIEPTP 480
Db 421 CYNAAVNVFPNDCTNKNLIWVEAGNFQGVNQFKAICSGQTIIRIDQKKGSKQIEPTP 480
Qy 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNWPMICAWLVK 540
Db 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNWPMICAWLVK 540
Qy 541 NGYSTWASCAKWKVPDSENWAEKVPPTPINLGSARSPTTTPKSTPLSQNYALTPL 600
Db 541 NGYSTWASCAKWKVPDSENWAEKVPPTPINLGSARSPTTTPKSTPLSQNYALTPL 600
Qy 601 ASLDELALPWPSTPNTFPVAGTAETQNTGAGSKACODGQLSPTWSIEEDLACFAGAE 660
Db 601 ASLDELALPWPSTPNTFPVAGTAETQNTGAGSKACODGQLSPTWSIEEDLACFAGAE 660
Qy 661 LKXDFSEPLNLD 672
Db 661 LKXDFSEPLNLD 672

RESULT 2
ID Q84365 PRELIMINARY; PRT; 721 AA.
AC Q84365;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nonstructural protein.
GN Name=NS1;
OS Murine minute virus (Murine parvovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MVM;
RA MEDLINE=83143341; PubMed=6298737;
RA Astell C.R.; Thomson M.; Merchinsk M.; Ward D.C.;
RT "The complete DNA sequence of minute virus of mice, an autonomous
RL parvovirus.";
RL Nucleic Acids Res. 11:999-1018(1983).
```

```

[2]
RN RP SEQUENCE FROM N.A.
RC STRAIN=MVM;
RX MEDLINE=86115415; PubMed=3502703;
RA Astell C.R.; Gardiner E.M.; Tattersall P.;
RT "DNA sequence of the lymphotropic variant of minute virus of mice,
RT MVM(1), and comparison with the DNA sequence of the fibrotropic
RT prototype strain.";
RL J. Virol. 57:656-669(1986).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=MVM;
RX MEDLINE=87061199; PubMed=3783817;
RA Morgan W.R.; Ward D.C.;
RT "Three splicing patterns are used to excise the small intron common to
RT all minute virus of mice RNAs.";
RL J. Virol. 60:1170-1174(1986).
DR EMBL; J02275; AAA67108.1; -
DR GO; GO:0019012; C:virion; IEA.
DR InterPro; IPR001257; P:Viral genome replication; IEA.
DR Pfam; PF01057; Parvo_Ns1; I.
KW Nonstructural protein.
SQ SEQUENCE 721 AA; 81896 MW; 18391758E42F0DCF CRC64;

Query Match      99.6%; Score 3623; DB 2; Length 721;
Best Local Similarity 99.7%; Pred. No. 7.3e-260;
Matches 670; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAGNAYSDEVLGATNMLKEKSNQEVFSFVKENNVQNGKDIGNWSYKKELODELKSLQ 60
Db 50 MAGNAYSDEVLGATNMLKEKSNQEVFSFVKENNVQNGKDIGNWSYKKELODELKSLQ 109
Qy 61 RGAETTWQSEDMEWETTVDEMCKQVFI FDSL VKKCLFEVLNTKNIFPGDVNMFVQHEW 120
Db 110 RGAETTWQSEDMEWETTVDEMCKQVFI FDSL VKKCLFEVLNTKNIFPGDVNMFVQHEW 169
Qy 121 GKQGWCHVLIIGKGFDSQAQGWRRQLNVLVYSLVLTACNVQLTPAERIKLREIAEDN 180
Db 170 GKQGWCHVLIIGKGFDSQAQGWRRQLNVLVYSLVLTACNVQLTPAERIKLREIAEDN 229
Qy 181 EWTLLTYKHQKQTKDYTCVLFQGNMAYFLTKKKISTSPRRGGYFLSSDSGWKTNFL 240
Db 230 EWTLLTYKHQKQTKDYTCVLFQGNMAYFLTKKKISTSPRRGGYFLSSDSGWKTNFL 289
Qy 241 KEGERHLVSKLYTDDMRPEVTETVTTAQTETKGRGRIOTKKEVSIKTTLKELVHKRVTSPE 300
Db 290 KEGERHLVSKLYTDDMRPEVTETVTTAQTETKGRGRIOTKKEVSIKTTLKELVHKRVTSPE 349
Qy 301 DMMWMPDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSPLD 360
Db 350 DMMWMPDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSPLD 409
Qy 361 TRACRIFAFHGNVYKCHAI CCVLNRQGGKRNVTLPFGPASTGKSI IAAIAQAVGNVG 420
Db 410 TRCIRIFAFHGNVYKCHAI CCVLNRQGGKRNVTLPFGPASTGKSI IAAIAQAVGNVG 469
Qy 421 CYNAAVNVFPNDCTNKNLIWVEAGNFQGVNQFKAICSGQTIIRIDQKKGSKQIEPTP 480
Db 470 CYNAAVNVFPNDCTNKNLIWVEAGNFQGVNQFKAICSGQTIIRIDQKKGSKQIEPTP 529
Qy 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNWPMICAWLVK 540
Db 530 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPDGLVDKNWPMICAWLVK 589
Qy 541 NGYSTWASCAKWKVPDSENWAEKVPPTPINLGSARSPTTTPKSTPLSQNYALTPL 600
Db 590 NGYSTWASCAKWKVPDSENWAEKVPPTPINLGSARSPTTTPKSTPLSQNYALTPL 649
Qy 601 ASLDELALPWPSTPNTFPVAGTAETQNTGAGSKACODGQLSPTWSIEEDLACFAGAE 660
Db 650 ASLDELALPWPSTPNTFPVAGTAETQNTGAGSKACODGQLSPTWSIEEDLACFAGAE 709
```

```
QY 661 LKXDFSEPLNLD 672
Db 710 LKXDFSEPLNLD 721

RESULT 3
VNC5 MUMIM STANDARD; PRT; 672 AA.
ID VNC5 MUMIM STANDARD; PRT; 672 AA.
AC P07300; P10837;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVP1).
GN Name=NS1;
OS Murine minute virus (strain MUM1) (Murine parvovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10795;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86115415; PubMed=3502703;
RA Astell C.R., Gardiner E.M., Tattersall P.;
RT "DNA sequence of the lymphotropic variant of minute virus of mice,
RT MVM(i), and comparison with the DNA sequence of the fibrotropic
RT prototype strain."
RL J. Virol. 57:656-669(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85242059; PubMed=3855242;
RA Sahli R., McMaster G.K., Hirt B.;
RT "DNA sequence comparison between two tissue-specific variants of the
RT autonomous parvovirus, minute virus of mice."
RL Nucleic Acids Res. 13:3617-3633(1985).
CC -1- FUNCTION: Seems necessary for viral DNA replication.
CC -1- SUBUNIT: Forms homooligomers. Interacts with human SYNCRIP (By
CC similarity).
CC -1- DOMAIN: The N-terminus (residues 1-275) possesses a negative effect
CC on transactivation (By similarity).
CC -1- DOMAIN: The C-terminus (residues 543-672) possesses an activation
CC domain (By similarity).
CC -1- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02481; -; NOT ANNOTATED CDS.
DR EMBL; M12032; AAA69567.1; -.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; 1.
KW ATP-binding; DNA replication; Noncapsid protein;
KW Nonstructural protein.
FT DOMAIN 1 276 Interacts with SYNCRIP (By similarity).
FT NP BIND 399 406 ATP (Potential).
FT FT CONFLICT 597 597 I -> L (in Ref. 2).
SQ SEQUENCE 672 AA; 76140 MW; 25F025FE328B4DF0 CRC64;

Query Match 97.7%; Score 3554; DB 1; Length 672;
Best Local Similarity 97.5%; Pred. No. 8.7e-255;
Matches 655; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVLGATNMLKESNEVFSFVKENNVQLNGKDIGNWSYKKEQLQDELKSLQ 60
Db 1 MAGNAYSDEVLGATNMLKESNEVFSFVKEDVQLNGKDIGNWNYSKKEQLQDELKSLQ 60
QY 61 RGAETTWQSDMEWETTVDEMTKKQVIFDLSLVKKCLFEVLNTKNIFPGDVNWFVQHEW 120
Db 61 RGAETTWQSDMEWETTVDEMTKKQVIFDLSLVKKCLFEVLNTKNIFADVTWVQHEW 120
QY 121 GKQGWCHVLIIGKQFSAQGWRRQLNYYWSRWLVTCNVQLTPAERIKLREIAEDN 180
```

RESULT 4

```
Q84363 PRELIMINARY; PRT; 721 AA.
ID Q84363
AC Q84363;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nonstructural protein.
GN Name=NS1;
OS Murine minute virus (Murine parvovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10794;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86115415; PubMed=3502703;
RA Astell C.R., Gardiner E.M., Tattersall P.;
RT "DNA sequence of the lymphotropic variant of minute virus of mice,
RT MVM(i), and comparison with the DNA sequence of the fibrotropic
RT prototype strain."
RL J. Virol. 57:656-669(1986).
DR EMBL; M12032; AAA69566.1; -.
DR PIR; A23008; UYPVIN.
DR GO; GO:0019012; C:virion; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; 1.
KW Nonstructural protein.
SQ SEQUENCE 721 AA; 81862 MW; 9FD293C27C7F4BBF CRC64;
```

Query Match 97.7%; Score 3554; DB 2; Length 721;
Best Local Similarity 97.5%; Pred. No. 9.6e-255;
Matches 655; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVLGATNWLKESNQEVFSFVKENVQNGKDIGWNSYKKELQDELKSLQ 60
 Db 50 MAGNAYSDEVLGATNWLKESNQEVFSFVKEDVQNGKDIGWNSYKKELQDELKSLQ 109
 QY 61 RGAETTWQSEDMEWETTVDKTKQVFI FDSLVKKCLFEVLNTKNIPFGDVNMFVQHEW 120
 Db 110 RGAETTWQSEDMEWETTVDKTKQVFI FDSLVKKCLFEVLNTKNIPADVTWVQHEW 169
 QY 121 GKQGWCHVLIIGKDFSOAGKQWRRQNLVYWSRLVTACNVQTPAERIKLREIAEDN 180
 Db 170 GKQGWCHVLIIGKDFSOAGKQWRRQNLVYWSRLVTACNVQTPAERIKLREIAEDS 229
 QY 181 EWTLLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPRDGGYFLSSDSGKTNFL 240
 Db 230 EWTLLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPRDGGYFLSSDSGKTNFL 289
 QY 241 KEGERHLVSKLYTDDMRPETVETTTTAQETKRGRIQTKKEVSIKTTLLKELVHKRVTSPE 300
 Db 290 KEGERHLVSKLYTDDMRPETVETTTTAQETKRGRIQTKKEVSIKTTLLKELVHKRVTSPE 349
 QY 301 DMMMQPDSYIEMMAQPGGENLLKNTLEICTLTLARTKTAFDLILEKAETSKLTFNLSLPD 360
 Db 350 DMMMQPDSYIEMMAQPGGENLLKNTLEICTLTLARTKTAFDLILEKAETSKLTFNLSLPD 409
 QY 361 TRACRIFAFHGMNYYKVCHAI CCVLNRQGGKRNVLPHGPASTGKSI IAAQIAQAVGNVG 420
 Db 410 TRTCKIFAFHGMNYYKVCHAI CCVLNRQGGKRNVLPHGPASTGKSI IAAQIAQAVGNVG 469
 QY 421 CYNAAVNFPPNDCTNKNLIWVEEAGNFGQVNOFKAI CSQTIRIDOKGKSGKQIEPTP 480
 Db 470 CYNAAVNFPPNDCTNKNLIWVEEAGNFGQVNOFKAI CSQTIRIDOKGKSGKQIEPTP 529
 QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNWPMICAWLVK 540
 Db 530 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNWPMICAWLVK 589
 QY 541 NGYSTWASYCAKWKGVDPWSNWAEPKVPPTPINLGSARSPTTPKSTPLSQNYALTPL 600
 Db 590 NGYSTWASYCAKWKGVDPWSNWAEPKVPPTPINLGSARSPTTPKSTPLSQNYALTPL 649
 QY 601 ASDLEDLALPWSPTNPVAGTAETQNTGEGSKACQDQGLSPTWSEIEDLRACFGAEP 660
 Db 650 ASDLEDLALPWSPTNPVAGTAETQNTGEGSKACQDQGLSPTWSEIEDLRACFGAEP 709
 QY 661 LKPDFSEPLNLD 672
 Db 710 LKPDFSEPLNLD 721

RESULT 5

Q83429 ID Q83429 PRELIMINARY; PRT; 672 AA.
 AC Q83429;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Nonstructural protein 1.
 GN Name=NS1;
 OS Mouse parvovirus 1.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
 OX NCBI_TaxID=35340;
 RN [1]
 RP MEDLINE=94365951; PubMed=8083985;
 RA Ball-Goodrich L.J., Johnson E.;
 RT "Molecular characterization of a newly recognized mouse parvovirus.";
 J. Virol. 68:6476-6486(1994).
 DR EMBL; U12469; AAA61405.1;
 DR GO; GO:0019012; C:virion; IEA.
 DR GO; GO:0019079; P:virion genome replication; IEA.
 DR InterPro; IPR001257; Parvo NS1.
 DR Pfam; PF01057; Parvo NS1; I.
 KW Nonstructural protein.

SQ SEQUENCE 672 AA; 76112 MW; 31C6365276727363 CRC64;
 Query Match 97.7%; Score 3552; DB 2; Length 672;
 Best Local Similarity 97.3%; Pred. No. 1.2e-254;
 Matches 654; Conservative 10; Mismatches 8; Indels 0; Gaps 0;
 QY 1 MAGNAYSDEVLGATNWLKESNQEVFSFVKENVQNGKDIGWNSYKKELQDELKSLQ 60
 Db 1 MAGNAYSDEVLGATNWLKESNQEVFSFVKEDVQNGKDIGWNSYKKELQDELKSLQ 60
 QY 61 RGAETTWQSEDMEWETTVDKTKQVFI FDSLVKKCLFEVLNTKNIPFGDVNMFVQHEW 120
 Db 61 RGAETTWQSEDMEWETTVDKTKQVFI FDSLVKKCLFEVLNTKNIPADVTWVQHEW 120
 QY 121 GKQGWCHVLIIGKDFSOAGKQWRRQNLVYWSRLVTACNVQTPAERIKLREIAEDN 180
 Db 121 GKQGWCHVLIIGKDFSOAGKQWRRQNLVYWSRLVTACNVQTPAERIKLREIAEDS 180
 QY 181 EWTLLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPRDGGYFLSSDSGKTNFL 240
 Db 181 EWTLLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPRDGGYFLSSDSGKTNFL 240
 QY 241 KEGERHLVSKLYTDDMRPETVETTTTAQETKRGRIQTKKEVSIKTTLLKELVHKRVTSPE 300
 Db 241 KEGERHLVSKLYTDDMRPETVETTTTAQETKRGRIQTKKEVSIKTTLLKELVHKRVTSPE 300
 QY 301 DMMMQPDSYIEMMAQPGGENLLKNTLEICTLTLARTKTAFDLILEKAETSKLTFNLSLPD 360
 Db 301 DMMMQPDSYIEMMAQPGGENLLKNTLEICTLTLARTKTAFDLILEKAETSKLTFNLSLPD 360
 QY 361 TRACRIFAFHGMNYYKVCHAI CCVLNRQGGKRNVLPHGPASTGKSI IAAQIAQAVGNVG 420
 Db 361 TRTCKIFAFHGMNYYKVCHAI CCVLNRQGGKRNVLPHGPASTGKSI IAAQIAQAVGNVG 420
 QY 421 CYNAAVNFPPNDCTNKNLIWVEEAGNFGQVNOFKAI CSQTIRIDOKGKSGKQIEPTP 480
 Db 421 CYNAAVNFPPNDCTNKNLIWVEEAGNFGQVNOFKAI CSQTIRIDOKGKSGKQIEPTP 480
 QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNWPMICAWLVK 540
 Db 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNWPMICAWLVK 540
 QY 541 NGYSTWASYCAKWKGVDPWSNWAEPKVPPTPINLGSARSPTTPKSTPLSQNYALTPL 600
 Db 541 NGYSTWASYCAKWKGVDPWSNWAEPKVPPTPINLGSARSPTTPKSTPLSQNYALTPL 600
 QY 601 ASDLEDLALPWSPTNPVAGTAETQNTGEGSKACQDQGLSPTWSEIEDLRACFGAEP 660
 Db 601 ASDLEDLALPWSPTNPVAGTAETQNTGEGSKACQDQGLSPTWSEIEDLRACFGAEP 660
 QY 661 LKPDFSEPLNLD 672
 Db 661 LKPDFSEPLNLD 672

RESULT 6

Q8JV14 ID Q8JV14 PRELIMINARY; PRT; 672 AA.
 AC Q8JV14;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Nonstructural protein 1.
 GN Name=NS1;
 OS Rat minute virus 1c.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
 OX NCBI_TaxID=172387;
 RN [1]
 RP MEDLINE=22120170; PubMed=12124471;
 RA Wan C.H., Soderlund-Venermo M., Pintel D.J., Riley L.K.;
 RT "Molecular characterization of three newly recognized rat parvoviruses."


```
RL J. Gen. Virol. 83:2075-2083 (2002).
RN [2]
RP Nonstructural protein 1.
RA Name=ns1;
RA Kilham rat virus.
RA Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF332884; AAM93279.1; -.
DR GO; GO:0019012; C:virion; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR001257; Parvo_Ns1.
DR Pfam; PF01057; Parvo_Ns1; 1.
KW Nonstructural protein.
SQ SEQUENCE 672 AA; 75988 MW; 52DF6549349CF3FD CRC64;

Query Match 92.0%; Score 3345; DB 2; Length 672;
Best Local Similarity 91.7%; Pred. No. 2.8e-239;
Matches 616; Conservative 22; Mismatches 34; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVLGATNWLKESQNOVFVFNENVQNLGKDIGNSYKKELODELKSLQ 60
DB 1 MAGNAYSDEVLGATNWLKSSQEVFVFNENVQNLGKDIGNSYKKELODELKSLQ 60
QY 61 RGAETTWQSEDMEWETTVDEMTHKQVIFDLSLVKKCLFEVLNTKNIFPGDVNMFVQHEW 120
DB 61 RGAETTWQSEDMEWESAVIDMTKKQVIFDLSLVKKCLFEVLSTKNIAPSDVTWVQHEW 120
QY 121 GKQGGHCHVLIGKQFSQAGKWRRLQNLVYRSRLVTCNVQLTPAERIKLREIAEDN 180
DB 121 GKQGGHCHVLIGKQFSQAGKWRRLQNLVYRSRLVTCNVQLTPAERIKLREIAEDS 180
QY 181 EWTLLTYKHQTKDYKCVLFGNMIAYFLTKKISTSPPRDGGYFLSSDGSWKTNFL 240
DB 181 EWTLLTYKHQTKDYKCVLFGNMIAYFLSKKICTSPPRDGGYFLSSDGSWKTNFL 240
QY 241 KEGERHLVSKLYTDMRPETVETVTTAQTAKRGRIQTKKEVSIKTTLKELVHKRVTSPE 300
DB 241 KEGERHLVSKLYTDMRPETVETVTTAQTAKRGRIQTKKEVSIKTTLKELVHKRVTSPE 300
QY 301 DWMQMOPDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360
DB 301 DWMQMOPDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSMAS 360
QY 361 TRACRIPAFHGMNKKVCHAIACVNLNROGGKRNVLPHGPASTGKSIIAQAIAGVNVG 420
DB 361 TRTCRIPAFHGMNKKVCHAIACVNLNROGGKRNVLPHGPASTGKSIIAQAIAGVNVG 420
QY 421 CYNAANVNFNDCTNKNLIWVEAGNFGQOVNQFKAIACSGQTIRIDQKKGSKQIEPTP 480
DB 421 CYNAANVNFNDCTNKNLIWVEAGNFGQOVNQFKAIACSGQTIRIDQKKGSKQIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
DB 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
QY 541 NGYQSTWASYCAKWKVPDMSENWAEKVPPTNLLGSARSPTTPKSTPLSQNYALTPL 600
DB 541 NGYQSTWASYCAKWKVPDMSENWAEKVPPTNLLGSARSPTTPKSTPLSQNYALTPL 600
QY 601 ASDLEDLALPWSPTNTPVAGTAETQNTGAGSKACODGOLSPTWSEIEDLACFCAEP 660
DB 601 ASDLADLALPWSPTNTPVAGTAASQNTGAGFTACQAGRSPTWSEIEDLACFQSEQ 660
QY 661 LKXDFSEPLNLD 672
DB 661 LKXDFSEPLNLD 672

RESULT 7
Q8JV28 PRELIMINARY; PRT; 672 AA.
ID Q8JV28
AC Q8JV28
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
```

```
Db 661 LKDFSEPLNLD 672
    |:||||:|
661 LEKDFSDSLTLD 672

RESULT 8
P88899 PRELIMINARY; PRT; 672 AA.
ID P88899
AC P88899;
DT 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Non-capsid protein.
GN Name=NS1;
OS Kilham rat virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=12441;
RN [1]
RP SEQUENCE FROM N.A.
RA Brown D.W., Like A.A.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U79033; AAB38326.1; -.
DR GO: 0019079; P: viral genome replication; IEA.
DR InterPro: IPR001257; Parvo_NSI.
DR Pfam; PF01057; Parvo_NSI; I.
SQ SEQUENCE 672 AA; 76056 MW; 0299B36871A0A10A CRC64;

Query Match 91.7%; Score 3335; DB 2; Length 672;
Best Local Similarity 91.4%; Pred. No. 1.5e-238;
Matches 614; Conservative 22; Mismatches 36; Indels 0; Gaps 0;

Qy 1 MAGNAYSDVLGATNWLKESNQVFSFVKNNVQLNGKDIGNSYKKELODELKSLQ 60
Db 1 MAGNAYSDVLGATNWLKSSQVFSFVKNNVQLNGKDIGNSYKKELODELKSLQ 60

Qy 61 RGAETTWQSDSEDMEWETTVDENTKKQVFI FDSL VKKCLFEVLNTKNI FPGDVNWFVQHEW 120
Db 61 RGAETTWQSDSEDMEWESA VDDMTKKQVFI FDSL VKKCLFEVLNTKNI FPGDVNWFVQHEW 120

Qy 121 GKDOGWCHVLIGGKDFSOAGKWRRLNVYWSRWLVTAACNVLTTPAERI KLRLEAEDN 180
Db 121 GKDOGWCHVLIGGKDFSOAGKWRRLNVYWSRWLVTAACNVLTTPAERI KLRLEAEDS 180

Qy 181 EWTLLTYKHQTKKDYTKCVLFGNMIAYFLTKKISTSPPRDGGYFLSSDGSWKTNFL 240
Db 181 EWTLLTYKHQTKKDYTKCVLFGNMIAYFLSKKICTSPPRDGGYFLSSDGSWKTNFL 240

Qy 241 KEGERHLVSKLYTDDMRPEVTETTTAQTGKRIQTKKEVSIKTTLKELVHKRVTSPE 300
Db 241 KEGERHLVSKLYTDDMRPEVTETTTAQAQKGRIGTRKEVSIKTTLKELVHKRVTSPE 300

Qy 301 DMMWQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360
Db 301 DMMWQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLANFSMAS 360

Qy 361 TRACRIFAFHGNVYKVCCHAI CCVLNRQGGKRNVL FHPGASTGKSIIAQAIAQVGNVG 420
Db 361 TRTCRIFAEHGNVYKVCCHAI CCVLNRQGGKRNVL FHPGASTGKSIIAQAIAQVGNVG 420

Qy 421 CYNAAVNFPNDCTNKNLIWVEEAGNFGQVNOQKAI CSQGTIRIDQKSGSKQIEPTP 480
Db 421 CYNAAVNFPNDCTNKNLIWVEEAGNFGQVNOQKAI CSQGTIRIDQKSGSKQIEPTP 480

Qy 601 ASLDELALBPWSTPNTFVAGTAETONTGAGSKACQDQGLSPTWSEIEEDLRACFGAEP 660
Db 601 ASLADLALBPWSTPNTFVAGTAASQNTGAGTACQGAQRSPTWSEIEADLRACFSQEQ 660

Qy 661 LKDFSEPLNLD 672
    |:||||:|
661 LEKDFSDSLTLD 672

RESULT 9
Q8JV18 PRELIMINARY; PRT; 672 AA.
ID Q8JV18
AC Q8JV18;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Nonstructural protein 1.
GN Name=NS1;
OS Rat minute virus 1a.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=172385;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=22120170; PubMed=12124471;
RA Wan C.H., Soderlund-Venermo M., Pintel D.J., Riley L.K.;
RT "Molecular Characterization of three newly recognized rat parvoviruses.";
RL J. Gen. Virol. 83:2075-2083 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Wan C.-H., Pintel D.J., Soderlund-Venermo M., Riley L.K.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF332882; AAM93275.1; -.
DR GO: 0019012; C: viralion; IEA.
DR GO: 0019079; P: viral genome replication; IEA.
DR InterPro: IPR001257; Parvo_NSI.
DR Pfam; PF01057; Parvo_NSI; I.
KW Nonstructural protein.
SQ SEQUENCE 672 AA; 76059 MW; 63D8B9EBF99E07B3 CRC64;

Query Match 91.5%; Score 3327; DB 2; Length 672;
Best Local Similarity 91.2%; Pred. No. 6e-239;
Matches 613; Conservative 23; Mismatches 36; Indels 0; Gaps 0;

Qy 1 MAGNAYSDVLGATNWLKESNQVFSFVKNNVQLNGKDIGNSYKKELODELKSLQ 60
Db 1 MAGNAYSDVLGATNWLKSSQVFSFVKNNVQLNGKDIGNSYKKELODELKSLQ 60

Qy 61 RGAETTWQSDSEDMEWETTVDENTKKQVFI FDSL VKKCLFEVLNTKNI FPGDVNWFVQHEW 120
Db 61 RGAETTWQSDSEDMEWESA VDDMTKKQVFI FDSL VKKCLFEVLNTKNI FPGDVNWFVQHEW 120

Qy 121 GKDOGWCHVLIGGKDFSOAGKWRRLNVYWSRWLVTAACNVLTTPAERI KLRLEAEDN 180
Db 121 GKDOGWCHVLIGGKDFSOAGKWRRLNVYWSRWLVTAACNVLTTPAERI KLRLEAEDS 180

Qy 181 EWTLLTYKHQTKKDYTKCVLFGNMIAYFLTKKISTSPPRDGGYFLSSDGSWKTNFL 240
Db 181 EWTLLTYKHQTKKDYTKCVLFGNMIAYFLSKKICTSPPRDGGYFLSSDGSWKTNFL 240

Qy 241 KEGERHLVSKLYTDDMRPEVTETTTAQTGKRIQTKKEVSIKTTLKELVHKRVTSPE 300
Db 241 KEGERHLVSKLYTDDMRPEVTETTTAQAQKGRIGTRKEVSIKTTLKELVHKRVTSPE 300

Qy 301 DMMWQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360
Db 301 DMMWQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLANFSMAN 360

Qy 361 TRACRIFAFHGNVYKVCCHAI CCVLNRQGGKRNVL FHPGASTGKSIIAQAIAQVGNVG 420
Db 361 TRTCRIFAEHGNVYKVCCHAI CCVLNRQGGKRNVL FHPGASTGKSIIAQAIAQVGNVG 420

Qy 421 CYNAAVNFPNDCTNKNLIWVEEAGNFGQVNOQKAI CSQGTIRIDQKSGSKQIEPTP 480
Db 421 CYNAAVNFPNDCTNKNLIWVEEAGNFGQVNOQKAI CSQGTIRIDQKSGSKQIEPTP 480
```

```
QY 481 VIMTTNITVVRIGCEERPEHTPIRDRMLNHLTHLPGDFGLVDKNEWPMICAWLVK 540
Db 481 VIMTTNITVVRIGCEERPEHTPIRDRMLNHLTHLPGDFGLVDKNEWPMICAWLVK 540
QY 541 NGYQSTWASYCAKWKVPDSEWNAEPKVPPTPINLLGSARSPTTPKSTPLSQNYALTPL 600
Db 541 NGYQSTWASYCAKWKVPDSEWNAEPKVPPTPINLLGSARSPTTPKSTPLSQNYALTPL 600
QY 601 ASDLELALPWPSTPNTPVAGTAETONTGEGSKACQDQGLSTWSEIEEDLRACFGAEP 660
Db 601 ASDLADLALPWPSTPNTPVAGTAASQNTGEGSPTWSEIEADLRACFSQEQ 660
QY 661 LKXDFSEPLNLD 672
Db 661 LKXDFSEPLNLD 672

RESULT 10
VNC5_PAVHH STANDARD; PRT; 672 AA.
AC P03133;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVP1)..
GN Name=NS1;
OS Hamster parvovirus H1.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10799;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83112183; PubMed=6823009;
RA Rhode S.L. III, Paradiso P.R.;
RT "Parvovirus genome: nucleotide sequence of H-1 and mapping of its
RT genes by hybrid-arrested translation.";
RL J. Virol. 45:173-184(1983).
CC -1- FUNCTION: Seems necessary for viral DNA replication.
CC -1- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X01457; CAA25689.1; -.
DR PIR; A03695; UYPPV1.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; 1.
KW ATP-binding; DNA replication; Noncapsid protein;
KW Nonstructural protein.
FT NP BIND 399 406 ATP (Potential).
SQ SEQUENCE 672 AA; 75993 MW; 12F331142F72AA6D CRC64;

Query Match 91.4%; Score 3326; DB 1; Length 672;
Best Local Similarity 91.1%; Pred. No. 7.2e-238;
Matches 612; Conservative 24; Mismatches 36; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVILGATNWLKESNQEVFSVFKNENVLQNGDKIGWNSYKKELODELKSLQ 60
Db 1 MAGNAYSDEVILGATNWLKSSQEVFSVFKNENVLQNGDKIGWNSYKKELODELKSLQ 60
QY 61 RGAETTDQSDMEDMEWETTYDQKQVFIQSLVKKCLFEVLNTKNTFPGDVNWFVQHEW 120
Db 61 RGAETTDQSDMEDMEWESAVDDMTKQVFIQSLVKKCLFEVLNTKNTAPSNVTWVQHEW 120
QY 121 GKQDGHCHVLIGKDFSQAGKWRRLQNVYWSRWLVTAQNVOLTPAERIKLREIAEDN 180
Db 121 GKQDGHCHVLIGKDFSQAGKWRRLQNVYWSRWLVTAQNVOLTPAERIKLREIAEDS 180
QY 181 EWTLLTYKHQTKKDYTKCVLFGNMIAYYFLSKKICTSPPRDGGYFLSSDSGWKTNFL 240
```

```
Db 181 EWTLLTYKHQTKKDYTKCVLFGNMIAYYFLSKKICTSPPRDGGYFLSSDSGWKTNFL 240
QY 241 KEGERHLVSKLYTDDMRPETVETVTTAQETKRGRIQTQKKEYSIKTKTLKELVHKRVTSPE 300
Db 241 KEGERHLVSKLYTDEMKEPETVETVTTAQEAKRGRIQTREEVSIKTKTLKELVHKRVTSPE 300
QY 301 DNMWQPDYSYIEMMAQPGGENLLKNTLCTLTARTKTAFDLLEKAETSCLTNFSLPD 360
Db 301 DNMWQPDYSYIEMMAQPGGENLLKNTLCTLTARTKTAFDLLEKAETSCLTNFSLPD 360
QY 361 TRACRIFAEGHWNVYKCHAI CCVLNRQGGKNTVLFHGPASTGKSIIAQAIAQAVGNVG 420
Db 361 TRTCRIFAEGHWNVYKCHAI CCVLNRQGGKNTVLFHGPASTGKSIIAQAIAQAVGNVG 420
QY 421 CYNAAANVPFNDCTNKNLIWVEEAGNFGQVQNFKAICSGQTIRIDQKGGKSGKQIEPTP 480
Db 421 CYNAAANVPFNDCTNKNLIWVEEAGNFGQVQNFKAICSGQTIRIDQKGGKSGKQIEPTP 480
QY 481 VIMTTNITVVRIGCEERPEHTPIRDRMLNHLTHLPGDFGLVDKNEWPMICAWLVK 540
Db 481 VIMTTNITVVRIGCEERPEHTPIRDRMLNHLTHLPGDFGLVDKNEWPMICAWLVK 540
QY 541 NGYQSTWASYCAKWKVPDSEWNAEPKVPPTPINLLGSARSPTTPKSTPLSQNYALTPL 600
Db 541 NGYQSTWASYCAKWKVPDSEWNAEPKVPPTPINLLGSARSPTTPKSTPLSQNYALTPL 600
QY 601 ASDLELALPWPSTPNTPVAGTAETONTGEGSKACQDQGLSTWSEIEEDLRACFGAEP 660
Db 601 ASDLADLALPWPSTPNTPVAGTAASQNTGEGSPTWSEIEADLRACFSQEQ 660
QY 661 LKXDFSEPLNLD 672
Db 661 LKXDFSEPLNLD 672

RESULT 11
Q8JVI6 PRELIMINARY; PRT; 672 AA.
AC Q8JVI6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nonstructural protein 1.
GN Name=NS1;
OS Rat minute virus 1b.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=172386;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22120170; PubMed=12124471;
RA Wan C.-H., Soderlund-Venermo M., Pintel D.J., Riley L.K.;
RT "Molecular characterization of three newly recognized rat
RT parvoviruses.";
RL J. Gen. Virol. 83:2075-2083(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Wan C.-H., Pintel D.J., Soderlund-Venermo M., Riley L.K.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF332883; AAM93277.1; -.
DR GO; GO:0019012; C:virion; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; 1.
KW Nonstructural protein.
SQ SEQUENCE 672 AA; 76201 MW; C2F1A71E6EF49A6 CRC64;

Query Match 91.4%; Score 3323; DB 2; Length 672;
Best Local Similarity 90.9%; Pred. No. 1.2e-237;
Matches 611; Conservative 24; Mismatches 37; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVILGATNWLKESNQEVFSVFKNENVLQNGDKIGWNSYKKELODELKSLQ 60
```

Db 1 MAGNAYSDEVILGATNWLKDSQEVSVFVFNENNVQNLGDKDIGNSVKRLQDELKSLQ 60
Qy 61 RGAETTTDQSDMEWETTVDMETTKQVFIPLSLVKKCLLFEVLNTKNIFFPGDVNMFVQHEW 120
Db 61 RGAETTTDQSDMEWESAVDDMTKQVFIPLSLVKKCLLFEVLSTKNIAPSDVTWVQHEW 120
Qy 121 GKDOGWCHVLIGKQFSQAQGWRRQLNLYVSRWLVTACNVLTTPAERIKLREIAEDN 180
Db 121 GKDOGWCHVLIGKQFSQAQGWRRQLNLYVSRWLVTACNVLTTPAERIKLREIAEDS 180
Qy 181 EWTLLTYKHQTKDYTKCVLFGNMIAYFELTKKISTSPRGGYFSLSDSGWKTNFL 240
Db 181 EWTLLTYKHQTKDYTKCVLFGNMIAYFELTKKISTSPRGGYFSLSDSGWKTNFL 240
Qy 241 KEGERHLVSKLYTDDMRPETVETTTAQTKEGRIOTKKEVSIKTKLVLHVKRVTSP 300
Db 241 KEGERHLVSKLYTDDMRPETVETTTAQTKEGRIOTKKEVSIKTKLVLHVKRVTSP 300
Qy 301 DMMNQDPSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLNFSLPD 360
Db 301 DMMNQDPSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLNFSLPD 360
Qy 361 TRACRIFAFHGMNVIKVCCHALCCVLNRQGGKNTVLFHGPASTGKSIIAQAIAGVGNV 420
Db 361 TRACRIFAFHGMNVIKVCCHALCCVLNRQGGKNTVLFHGPASTGKSIIAQAIAGVGNV 420
Qy 421 CYNAAVNFPPNDCTNKNLIWVEEAGNFGQVNOFKAI CSQTIRIDQKKGSKQIEPTP 480
Db 421 CYNAAVNFPPNDCTNKNLIWVEEAGNFGQVNOFKAI CSQTIRIDQKKGSKQIEPTP 480
Qy 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVK 540
Db 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVK 540
Qy 541 NGYOSTWASYCAKWKGVDPDSEDAEPKLETPINSLGMSRSPSTPSTPLSNQYALTPL 600
Db 541 NGYOSTWASYCAKWKGVDPDSEDAEPKLETPINSLGMSRSPSTPSTPLSNQYALTPL 600
Qy 601 ASDLEDLALPWSPTNPVAGTAETQNTGEGAGSKACQDGLSPTWSIEEDLACFGAEP 660
Db 601 ASDLEDLALPWSPTNPVAGTAETQNTGEGAGSKACQDGLSPTWSIEEDLACFGAEP 660
Qy 661 LKDDFSEPLNLD 672
Db 661 LKDDFSEPLNLD 672

RESULT 12

VNCS_PAVL3 STANDARD; PRT; 668 AA.
AC P36311;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 05-JUN-2004 (Rel. 44, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1).
GN Name=NS1;
OS Parvovirus LuIII.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=35339;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93297126; PubMed=8517025;
RA Difford N., Chen K.C., Bates R.C., Lederma M.;
RT "The complete nucleotide sequence of parvovirus LuIII and localization
of a unique sequence possibly responsible for its encapsidation
pattern.";
RL Virology 192:339-345 (1993).
CC -!- FUNCTION: Seems necessary for viral DNA replication.
CC -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M81888; -; NOT_ANNOTATED_CDS.
DR PIR; A44276; A44276.
DR InterPro; IPR001257; Parvo_NS1.
DR Pfam; PF01057; Parvo_NS1; I.
KW ATP-binding; DNA replication; Noncapsid protein;
KW Nonstructural protein.
FT NP_BIND 399 406 ATP (Potential).
SQ SEQUENCE 668 AA; 75846 MW; CAB69049F8F86B53 CRC64;

Query Match 91.3%; Score 3322; DB 1; Length 668;

Best Local Similarity 91.4%; Pred. No. 1.4e-237;

Matches 614; Conservative 23; Mismatches 31; Indels 4; Gaps 2;

Qy 1 MAGNAYSDEVILGATNWLKDSQEVSVFVFNENNVQNLGDKDIGNSVKRLQDELKSLQ 60

Db 1 MAGNAYSDEVILGATNWLKDSQEVSVFVFNENNVQNLGDKDIGNSVKRLQDELKSLQ 60

Qy 61 RGAETTTDQSDMEWETTVDMETTKQVFIPLSLVKKCLLFEVLNTKNIFFPGDVNMFVQHEW 120

Db 61 RGAETTTDQSDMEWESAVDDMTKQVFIPLSLVKKCLLFEVLSTKNIAPSDVTWVQHEW 120

Qy 121 GKDOGWCHVLIGKQFSQAQGWRRQLNLYVSRWLVTACNVLTTPAERIKLREIAEDN 180

Db 121 GKDOGWCHVLIGKQFSQAQGWRRQLNLYVSRWLVTACNVLTTPAERIKLREIAEDQ 180

Qy 181 EWTLLTYKHQTKDYTKCVLFGNMIAYFELTKKISTSPRGGYFSLSDSGWKTNFL 240

Db 181 EWTLLTYKHQTKDYTKCVLFGNMIAYFELTKKISTSPRGGYFSLSDSGWKTNFL 240

Qy 241 KEGERHLVSKLYTDDMRPETVETTTAQTKEGRIOTKKEVSIKTKLVLHVKRVTSP 300

Db 241 KEGERHLVSKLYTDDMRPETVETTTAQTKEGRIOTKKEVSIKTKLVLHVKRVTSP 300

Qy 301 DMMNQDPSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLNFSLPD 360

Db 301 DMMNQDPSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLNFSLPD 360

Qy 361 TRACRIFAFHGMNVIKVCCHALCCVLNRQGGKNTVLFHGPASTGKSIIAQAIAGVGNV 420

Db 361 TRACRIFAFHGMNVIKVCCHALCCVLNRQGGKNTVLFHGPASTGKSIIAQAIAGVGNV 420

Qy 421 CYNAAVNFPPNDCTNKNLIWVEEAGNFGQVNOFKAI CSQTIRIDQKKGSKQIEPTP 480

Db 421 CYNAAVNFPPNDCTNKNLIWVEEAGNFGQVNOFKAI CSQTIRIDQKKGSKQIEPTP 480

Qy 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVK 540

Db 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVK 540

Qy 541 NGYOSTWASYCAKWKGVDPDSEDAEPKLETPINSLGMSRSPSTPSTPLSNQYALTPL 600

Db 541 NGYOSTWASYCAKWKGVDPDSEDAEPKLETPINSLGMSRSPSTPSTPLSNQYALTPL 600

Qy 601 ASDLEDLALPWSPTNPVAGTAETQNTGEGAGSKACQDGLSPTWSIEEDLACFGAEP 660

Db 598 -SDLEDLALPWSPTNPVAGTAETQNTGEGAGSKACQDGLSPTWSIEEDLACFGAEP 656

Qy 661 LKDDFSEPLNLD 672

Db 657 WKSDESQPLNLD 668

RESULT 13

O71159
ID O71159 PRELIMINARY; PRT; 665 AA.
AC O71159;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)

```
01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nonstructural protein (fragment).
GN Name=NS1;
OS Kilham rat virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=12441;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=U-Mass;
RX MEDLINE=98184569; PubMed=9525656;
RA Ball-Goodrich L.J., Leland S.E., Johnson E.A., Paturzo F.X.,
RA Jacoby R.O.;
RT "Rat parvovirus type 1: the prototype for a new rodent parvovirus
RT serogroup.";
RL J. Virol. 72:3289-3299(1998).
DR EMBL; AF036711; AAC40695.1; -.
DR GO; GO:0019012; C:virion; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR001257; Parvo_Ns1.
DR Pfam; PF01057; Parvo_Ns1; 1.
DR Nonstructural protein.
KW Nonstructural protein.
FT NON TER 1
SQ SEQUENCE 665 AA; 75375 MW; 778E29043417E409 CRC64;

Query Match 90.8%; Score 3303; DB 2; Length 665;
Best Local Similarity 91.4%; Pred. No. 3.6e-236;
Matches 608; Conservative 22; Mismatches 35; Indels 0; Gaps 0;

Qy 8 DEVLGATNLKEKSNQVFSVPFKNVQNLGKDIGNSYKKELOBELKSLQGAETTW 67
Db 1 DEVLGATNLKDSQVFSVPFKNVQNLGKDIGNSYKKELOBELKSLQGAETTW 60

Qy 68 DQSEDMETVDEMTKKQVIFPSLVKKCLFEVLNTKNIFFPGVNNFVQHEWGDQGW 127
Db 61 DQSEDMESAVDMTKQVIFPSLVKKCLFEVLNTKNIAPSDVTFVQHEWGDQGW 120

Qy 128 CHVLIGKDFSOAGKWRRLQNTYWSRWLTACNVQLTPAERIKLREIADNFWTLLT 187
Db 121 CHVLIGKDFSOQGWRRRLQNTYWSRWLTACNVQLTPAERIKLREIADNFWTLLT 180

Qy 188 YKHQTKDYTKCVLFGNMIAYFLTKKISTSPRGGYFLSSDSGKTNFLKEGRHL 247
Db 181 YKHGHTKDYTKCVLFGNMIAYFLSKKICTSPRGGYFLSSDSGKTNFLKEGRHL 240

Qy 248 VSKLYTDMRPETVETTTAAQETKRGRIOTKKEVSIKTLKELVHKRVTSPEDDMMWQP 307
Db 241 VSKLYTDMRPETVETTTAAQETKRGRIOTKKEVSIKTLKELVHKRVTSPEDDMMWQP 300

Qy 308 DSYEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPDTRACRIF 367
Db 301 DSYEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPDTRACRIF 360

Qy 368 AFHGWNVYKCHAIACCVLNQGGKNTVLFHGPASTGKSIITAAQIAQAVGNVCYNA 427
Db 361 AEHGWNVYKCHAIACCVLNQGGKNTVLFHGPASTGKSIITAAQIAQAVGNVCYNA 420

Qy 428 NFPPNDCTNKNLIWVEBAGNFQGVNQPKAICSGQTIIDQKGSQIEPTPVMITNE 487
Db 421 NFPPNDCTNKNLIWVEBAGNFQGVNQPKAICSGQTIIDQKGSQIEPTPVMITNE 480

Qy 488 NITVVRIGCEPERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVKNGYSTM 547
Db 481 NITVVRIGCEPERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVKNGYSTM 540

Qy 548 ASYCAKWKGVDPDWSNNAEPKVPPIINLLGSARSPFTTPKSTPLSQNYALTPLASDLE 607
Db 541 ACYCAKWKGVDPDWSNNAEPKVPPIINLLGSARSPFTTPKSTPLSQNYALTPLASDLE 600

Qy 608 ALEPWSPTNTVAGTAETQNTGEAGSKACQDQSLSPTWSETEEDLACFGAEPLKDFSE 667
Db 601 ALEPWSPTNTVAGTAETQNTGEAGSKACQDQSLSPTWSETEEDLACFGAEPLKDFSE 660

Qy 668 PLNLD 672

Db 661 ELTLD 665

RESULT 14
071157 PRELIMINARY; PRT; 671 AA.
ID 071157
AC 071157
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nonstructural protein.
GN Name=NS1;
OS Rat parvovirus la.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=74581;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98184569; PubMed=9525656;
RA Ball-Goodrich L.J., Leland S.E., Johnson E.A., Paturzo F.X.,
RA Jacoby R.O.;
RT "Rat parvovirus type 1: the prototype for a new rodent parvovirus
RT serogroup.";
RL J. Virol. 72:3289-3299(1998).
DR EMBL; AF036710; AAC40693.1; -.
DR GO; GO:0019012; C:virion; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR001257; Parvo_Ns1.
DR Pfam; PF01057; Parvo_Ns1; 1.
DR Nonstructural protein.
KW Nonstructural protein.
SQ SEQUENCE 671 AA; 75752 MW; 9BCB39A39298D4DE CRC64;

Query Match 80.7%; Score 2935.5; DB 2; Length 671;
Best Local Similarity 81.1%; Pred. No. 6.8e-209;
Matches 545; Conservative 42; Mismatches 84; Indels 1; Gaps 1;

Qy 1 MAGNAYSEVLGATNLKEKSNQVFSVPFKNVQNLGKDIGNSYKKELOBELKSLQ 60
Db 1 MAGNAYSEVLGATNLKEKSNQVFSVPFKNVQNLGKDIGNSYKKELOBELKSLQ 60

Qy 61 RGABTTWDQSEDMETVDEMTKKQVIFPSLVKKCLFEVLNTKNIFFPGVNNFVQHEW 120
Db 61 RGABTTWDQSEDMETVDEMTKKQVIFPSLVKKCLFEVLNTKNIFFPGVNNFVQHEW 120

Qy 121 GKDQGHCHVLIGKDFSOAGKWRRLQNTYWSRWLTACNVQLTPAERIKLREIADN 180
Db 121 GKDQGHCHVLIGKDFSOAGKWRRLQNTYWSRWLTACNVQLTPAERIKLREIADN 180

Qy 181 EWTLLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPRGGYFLSSDSGKTNFL 240
Db 181 EWTLLTYKHQTKDYTKCVLFGNMIAYFLTKKISTSPRGGYFLSSDSGKTNFL 240

Qy 241 KEGERHLVSKLYTDMRPETVETTTAAQETKRGRIOTKKEVSIKTLKELVHKRVTSPE 300
Db 241 KEGERHLVSKLYTDMRPETVETTTAAQETKRGRIOTKKEVSIKTLKELVHKRVTSPE 300

Qy 301 DMMWQPDSEYEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360
Db 301 DMMWQPDSEYEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360

Qy 361 TRACRIIFAHGWNVYKCHAIACCVLNQGGKNTVLFHGPASTGKSIITAAQIAQAVGNV 420
Db 361 TRACRIIFAHGWNVYKCHAIACCVLNQGGKNTVLFHGPASTGKSIITAAQIAQAVGNV 420

Qy 421 CYNAAVNFPPNDCTNKNLIWVEBAGNFQGVNQPKAICSGQTIIDQKGSQIEPTP 480
Db 421 CYNAAVNFPPNDCTNKNLIWVEBAGNFQGVNQPKAICSGQTIIDQKGSQIEPTP 480

Qy 481 VIMTTNENITVVRIGCEPERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
Db 481 VIMTTNENITVVRIGCEPERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540

Qy 541 NGYQSTWASCAKWKGVDPDWSNNAEPKVPPIINLLGSARSPFTTPKSTPLSQNYALTPL 600
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2005, 23:46:57 ; Search time 8541.75 Seconds
(without alignments)
11177.792 Million cell updates/sec

Title: US-10-069-056-12
Perfect score: 2019
Sequence: 1 atggctggaatgcttactc.....agccgtgaacttgactaa 2019

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2019	100.0	2019	6	AX137747 Sequence
2	2017.4	99.9	2019	6	AX137736 Sequence
3	2017.4	99.9	5081	14	V01115 Minute viru
4	2017.4	99.9	5149	14	J02275 Minute viru
5	2015.8	99.8	2019	6	AX137739 Sequence
6	2015.8	99.8	2019	6	AX137743 Sequence
7	2015.8	99.8	2019	6	AX137751 Sequence
8	1905.4	94.4	5085	14	MVU34253
9	1902.2	94.2	4764	14	MVU34253
10	1902.2	94.2	5087	14	PAMVMI
11	1899	94.1	5144	14	MPU12469
12	1873	92.8	4764	14	MOU34254
13	1862.2	92.2	4773	14	HOU34255
14	1855.8	91.9	4761	14	MVU34256
15	1771	87.7	5121	6	CQ786765
16	1662.2	82.3	4904	14	AF321230
17	1659	82.2	4795	14	AF322884
18	1657.4	82.1	4927	14	KRU79033
19	1656.6	82.1	5135	14	PVRSEQ

20	1655.8	82.0	4816	14	AF332883	AF332883 Rat minut
21	1652.6	81.9	4813	14	AF332882	AF332882 Rat minut
22	1644.6	81.5	5176	14	PARH1	X01457 Parvovirus
23	1644.4	81.4	3995	14	AF036711	AF036711 Kilham ra
24	1275	63.2	4936	14	AF036710	AF036710 Rat parvo
25	1051.6	52.1	2007	14	AB000062	AB000062 Feline pa
26	1050	52.0	2007	14	AB000048	AB000048 Feline pa
27	1050	52.0	2007	14	AB000057	AB000057 Feline pa
28	1050	51.8	5124	14	PVPFVC	M38246 Feline panl
29	1046.8	51.8	2007	14	AB000049	AB000049 Feline pa
30	1046.8	51.8	2007	14	AB000058	AB000058 Feline pa
31	1046.8	51.8	2007	14	AB000060	AB000060 Feline pa
32	1046.8	51.8	2007	14	AB000069	AB000069 Feline pa
33	1045.2	51.8	2007	14	AB000053	AB000053 Feline pa
34	1045.2	51.8	2007	14	AB000063	AB000063 Feline pa
35	1045.2	51.8	2007	14	AB000065	AB000065 Feline pa
36	1045.2	51.8	2007	14	AB000067	AB000067 Feline pa
37	1045.2	51.8	5075	14	PVCY1A	D26079 Canine parv
38	1043.6	51.7	2007	14	AB000051	AB000051 Feline pa
39	1043.6	51.7	4983	14	PAPFVNS1	X55115 Feline panl
40	1043.6	51.7	5124	14	PVCPVC	M38245 Canine parv
41	1042	51.6	2007	14	AB000055	AB000055 Feline pa
42	1042	51.6	5323	14	PVCCPN	M19296 Canine parv
43	1040.4	51.5	5049	6	AR043629	AR043629 Sequence
44	1040.4	51.5	5049	6	AR043630	AR043630 Sequence
45	1040.4	51.5	5094	14	PVMRFD	D00765 Mink enteri

ALIGNMENTS

RESULT 1	AX137747	AX137747	2019 bp	DNA	linear	PAT 30-MAY-2001
LOCUS	Sequence	12 from Patent	EP1077260.			
DEFINITION	Sequence	12 from Patent	EP1077260.			
ACCESSION	AX137747					
VERSION	AX137747.1	GI:14273921				
KEYWORDS	Mice minute virus					
SOURCE	Mice minute virus					
ORGANISM	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.					
REFERENCE	Nuesch, J. and Rommelaere, J.					
AUTHORS	Parvovirus ns1 variants					
TITLE	Patent: EP 1077260-A 12 21-FEB-2001;					
JOURNAL	Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts					
FEATURES	(DE)					
source	Location/Qualifiers					
	1..2019					
	/organism="Mice minute virus"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:10794"					
	/notes="Parvovirus NS1 variant"					
	1..2019					
	/notes="unnamed protein product"					
	/codon_start=1					
	/protein_id="CAC39995.1"					
	/db_xref="GI:14273921"					
	/translation="MAGNAYSDEVLTGATNWLKESNOEVFSFVKENVNOLNGKDIGM					
	NSYKELQDELKSLQRGATHTDQSEDMEWETTVDMTKKQVFIQDLSLVKKCLFEVL					
	NTKNTIPGVDNPNFVHEGWGQDGHCHVIGKDFSOAQGWRRQLNVTWSRWLVLA					
	CNVQLPARIKLEIAEDNWTLLTYHKQTKDYTCVLFNNIAYVFLTKKIS					
	TSPRRGGYFLSDSGKTNFLKGERHLVSLKLYTDDMRPTVETVTTVTAQTKRGI					
	OTKKVYSIKTTLKELVHKVTSPEDEMMWQPSYIEMMAQGGENLLKNTLEICTLFL					
	ARTKTAFLILEKATSKLTNPSLPTDTCRIEAFHGWYKVCYKICCVLNRQGGKR					
	NAVLFHGASTGKSIIAQALAAQAVNGVCYNAANVFPDNCNKLLIWEAAGNFQ					
	QVNFKAICSGGTIRIDQKGSQKQIEPTVIMTNITVVRIGCEERPEHTOPTKD					
	RLMNLHITLPLGDFGLGNWPMI CAWLKNGYQSTWASACAKMGKPDWSENWAE					
	PKVPTINILGARSPTTPKSTPLSQNTALPLASDLLEDALEPSTNTPTVAGTAE					
	TQNTGAGSKACQDGLSFTWSIEEDLRACFGAEPLKXDFSEPLNLD"					

ORIGIN

Query Match 100.0%; Score 2019; DB 6; Length 2019;

Best Local Similarity 100.0%; Pred. No. 0;			
Matches 2019; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	ATGCTGGAATGCTTACTCTGATCAAGCTTTTGGAGCAACCAACTGGTTAAAGGAAAA	60
Ds	1	ATGCTGGAATGCTTACTCTGATGAAGTTTGGAGCAACCAACTGGTTAAAGGAAAA	60
Qy	61	AGTAACCAAGGAGTGTCTCTATTTGTTTTTAAAAATGAAATGTTCAACTGAAATGGAATA	120
Ds	61	AGTAACCAAGGAGTGTCTCTATTTGTTTTTAAAAATGAAATGTTCAACTGAAATGGAATA	120
Qy	121	GATATCGATGGAATAGTTTACAAAAAGAGCTGAGAGGAGCGAGCTGAAATCTTTTCAA	180
Ds	121	GATATCGATGGAATAGTTTACAAAAAGAGCTGAGAGGAGCGAGCTGAAATCTTTTCAA	180
Qy	181	CGAGAGCGGAAACTTACTTGGGACCAAGCAGGACATGGAATGGGAAACACACAGTGGAT	240
Ds	181	CGAGAGCGGAAACTTACTTGGGACCAAGCAGGAGCATGGAAATGGGAAACACACAGTGGAT	240
Qy	241	GAATGACCAAAAAAGCAAGTATTCATTTTGTGTTTAAATAATGTTTATTTGAA	300
Ds	241	GAATGACCAAAAAAGCAAGTATTCATTTTGTGTTTAAATAATGTTTATTTGAA	300
Qy	301	GTGCTTAAACAAAGATATATTTCTGCTGATGTTAAATGGTTTGTGCAACATGAATGG	360
Ds	301	GTGCTTAAACAAAGATATATTTCTGCTGATGTTAAATGGTTTGTGCAACATGAATGG	360
Qy	361	GGAAAGACCAAGCTGCACTGCGCATGTAATTCGAGGAAAGGACTTTAGTCAAGCT	420
Ds	361	GGAAAGACCAAGCTGCACTGCGCATGTAATTCGAGGAAAGGACTTTAGTCAAGCT	420
Qy	421	CAAGGAAATGGTGGAGAGCAACTAAATGTTTACTGGAGCAGATGGTTGTAACAGCC	480
Ds	421	CAAGGAAATGGTGGAGAGCAACTAAATGTTTACTGGAGCAGATGGTTGTAACAGCC	480
Qy	481	TGTAATGCAACTAAACACAGCTGAAAGAAATTAACATAGAGAAATAGCAGAGCAAT	540
Ds	481	TGTAATGCAACTAAACACAGCTGAAAGAAATTAACATAGAGAAATAGCAGAGCAAT	540
Qy	541	GAGTGGGTACTCTACTTACTTATAAGCATAAGCAAAACCAAAAGACTATACCAAGTGT	600
Ds	541	GAGTGGGTACTCTACTTACTTATAAGCATAAGCAAAACCAAAAGACTATACCAAGTGT	600
Qy	601	GTTCTTTTGGAAACATGATGTTACTATTTTTTAACTAAAGAAAAATAGCACTAGT	660
Ds	601	GTTCTTTTGGAAACATGATGTTACTATTTTTTAACTAAAGAAAAATAGCACTAGT	660
Qy	661	CCACCAAGAGACGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAACTAATCTTTTA	720
Ds	661	CCACCAAGAGACGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAACTAATCTTTTA	720
Qy	721	AAAGAAGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG	780
Ds	721	AAAGAAGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG	780
Qy	781	GTTGAAACCAAGTAAACCTGCGCAGGAACTAAGCGCGCAGAAATTTCAAACTAAAAA	840
Ds	781	GTTGAAACCAAGTAAACCTGCGCAGGAACTAAGCGCGCAGAAATTTCAAACTAAAAA	840
Qy	841	GAAGTTTCTATTAAAACTACACTTAAAGAGCTGGTGCATAAAAAGAGTAACCTCACCAG	900
Ds	841	GAAGTTTCTATTAAAACTACACTTAAAGAGCTGGTGCATAAAAAGAGTAACCTCACCAG	900
Qy	901	GACTGGATGATGAGCAGCAGAGTTTACATTGAAATGATGGCTCAACCAAGTGGAGAA	960
Ds	901	GACTGGATGATGAGCAGCAGAGTTTACATTGAAATGATGGCTCAACCAAGTGGAGAA	960
Qy	961	AACCTGCTGAAAAATACGCTAGAGATTTGTACACTTAACCTAGCAGAAACCAAAACAGCA	1020
Ds	961	AACCTGCTGAAAAATACGCTAGAGATTTGTACACTTAACCTAGCAGAAACCAAAACAGCA	1020
Qy	1021	TTTGACTTAAATTTAGAAAAAGCTGAAACCAAGCACTAACCACCTTTTCACTGCCTGAC	1080

Ds	1021	TTTGACTTAAATTTAGAAAAAGCTGAAACCAAGCAAACTAACCAACTTTTCACTGCCTGAC	1080
Qy	1081	ACAAGAACCTGCAAGATTTTGTGTTTTCATGCGTGGAACTATGTTAAAGTTTGCATGCT	1140
Ds	1081	ACAAGAACCTGCAAGATTTTGTGTTTTCATGCGTGGAACTATGTTAAAGTTTGCATGCT	1140
Qy	1141	ATTGCTGTGTTTTTAAACAGACAGGAGGCAAAAGAAATCTGTTTTTATTTTCAATGGAACA	1200
Ds	1141	ATTGCTGTGTTTTTAAACAGACAGGAGGCAAAAGAAATCTGTTTTTATTTTCAATGGAACA	1200
Qy	1201	GCAGACACAGGCAAAATCTATTATGCAAGCCATAGCAAGAGTTGGCAATGTTGGT	1260
Ds	1201	GCAGACACAGGCAAAATCTATTATGCAAGCCATAGCAAGAGTTGGCAATGTTGGT	1260
Qy	1261	TGCTATAATGCAAGCAATGTAATCTTCCATTTTAAAGTGTACCAACAGAACTTGATT	1320
Ds	1261	TGCTATAATGCAAGCAATGTAATCTTCCATTTTAAAGTGTACCAACAGAACTTGATT	1320
Qy	1321	TGGGTAGAAGAACTGCTGTAATTTGGACAGCAAGTAAACCAAGTTTAAAGCATTGCTCT	1380
Ds	1321	TGGGTAGAAGAACTGCTGTAATTTGGACAGCAAGTAAACCAAGTTTAAAGCATTGCTCT	1380
Qy	1381	GGTCAAACTATTGCGATTTGATCAAAAGGAAAGGCAAGCAAGATTTGAACCAACCA	1440
Ds	1381	GGTCAAACTATTGCGATTTGATCAAAAGGAAAGGCAAGCAAGATTTGAACCAACCA	1440
Qy	1441	GTCAATCATGACCAAAATGAGAACATTTACAGTGGTGCAGATAGGCTGCGAAGAAAGACA	1500
Ds	1441	GTCAATCATGACCAAAATGAGAACATTTACAGTGGTGCAGATAGGCTGCGAAGAAAGACA	1500
Qy	1501	GAAACACACTCAACCAATCAGAGACAGAAATGCTTTAACTTATCATCAACACATCTTGCT	1560
Ds	1501	GAAACACACTCAACCAATCAGAGACAGAAATGCTTTAACTTATCATCAACACATCTTGCT	1560
Qy	1561	GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGCTGGTTGGTAAAG	1620
Ds	1561	GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGCTGGTTGGTAAAG	1620
Qy	1621	AATGGTTTACCAATCTACCATGGCAAGCTACTGTCTTAAATGGGGCAAAAGTTCTGATTGG	1680
Ds	1621	AATGGTTTACCAATCTACCATGGCAAGCTACTGTCTTAAATGGGGCAAAAGTTCTGATTGG	1680
Qy	1681	TCAGAAAACTGGCGGAGCCAAAGTTGCCAACTCTTATAAATTTAATAGTTTGGCAAGC	1740
Ds	1681	TCAGAAAACTGGCGGAGCCAAAGTTGCCAACTCTTATAAATTTAATAGTTTGGCAAGC	1740
Qy	1741	TCACCAATTCAGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTTAATCCACTT	1800
Ds	1741	TCACCAATTCAGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTTAATCCACTT	1800
Qy	1801	GCAATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAATATCTCTGTTGG	1860
Ds	1801	GCAATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAATATCTCTGTTGG	1860
Qy	1861	GGCACTGCGAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGTCAA	1920
Ds	1861	GGCACTGCGAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGTCAA	1920
Qy	1921	CTGAGCCCAACTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGTTCGTCGGGAACCG	1980
Ds	1921	CTGAGCCCAACTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGTTCGTCGGGAACCG	1980
Qy	1981	TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019	
Ds	1981	TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019	

RESULT 2
AX137736
LOCUS
DEFINITION
ACCESSION
VERSION

AX137736 2019 bp DNA linear PAT 30-MAY-2001
Sequence 1 from Patent EP1077260.
AX137736
AX137736.1 GI:14273909

KEYWORDS	Mice minute virus
SOURCE	Mice minute virus
ORGANISM	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE	1. Nueesch, J. and Rommelaere, J.
AUTHORS	Parvovirus ns1 variants
TITLE	Patent: EP 1077260-A 1 21-FEB-2001;
JOURNAL	Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts
(DE)	
FEATURES	Location/Qualifiers
source	1. .2019
	/organism="Mice minute virus"
	/mol_type="unassigned DNA"
	/db_xref="taxon:10794"
	/note="Wildtype Parvovirus NS1"
CDS	1. .2019
	/note="unnamed protein product"
	/codon_start=1
	/protein_id="CAC39989.1"
	/db_xref="GI:14273910"
	/translation="MAGNAYSDEVLAGTNLWLEKESNOEVSFVFKNENVLNGKIDIGW
	NSYKLEQDELKSLORGAEETWQSEDMEWETTVDKTKQVFIIDSLVKKCLPEVL
	NTKNIPGDVNNFVQHEWKDQGHCHVLIGKDFQAQGGKWRQLNLYVWSRLVTA
	CNVQLTPAEKIKUREAEDNEWVTLTKYKHKQKDYTKCVLFENMIAYFLFKKIS
	TSPPRDGVFLSSDGMKTNFLKEGHRHLVSKDYDDMPETVETVTTAQETKGR
	QTKVEYSIKTLLKELVHKRVTSPEWMMQPDYSIEMMAQPGGENLLKNTLEICTLTL
	ARTKAPDLLEKAETSKLNFSLPDRTRCRIFAFGWNVYKCHAI CCVLNRQGGKR
	NTVLFGPASTGKSIIAQATAOAGVNGVCYNAANVFPENDCTKNKLIWVEEAGNRQ
	QVNOFKAI CSQVIRIDQKSGSKQIEPTVIMTNENITVVRIGEEPEHTQPIRD
	RLNLIHLTHLPDGLVDKNEMPMICAMLVNGYQSTWASYCAKNGKVPDWSENAE
	KPYPTINLGSARSFPTPKSPPLSONYALTPLEADLEDALEPWSPTNPVAGTAE
	TQNTGAGSKACQDGLSPTWSBIEEDLRACFGAEPLKKDFSEPLND"
ORIGIN	
Query Match	99.9%; Score 2017.4; DB 6; Length 2019;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 2019; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
QY	1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGACCAACCACTGGTTTAAAGGAAAA 60
DB	1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGACCAACCACTGGTTTAAAGGAAAA 60
QY	61 AGTAACGAGGAAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGGAATGGA 120
DB	61 AGTAACGAGGAAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGGAATGGA 120
QY	121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCGAGGACGAGCTGAAATCTTTACA 180
DB	121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCGAGGACGAGCTGAAATCTTTACA 180
QY	181 CGAGGAGCGGAACTACTTGGGACCAAGCGAGGACATGGAATGGGAAACACACGTGG 240
DB	181 CGAGGAGCGGAACTACTTGGGACCAAGCGAGGACATGGAATGGGAAACACACGTGG 240
QY	241 GAAATGACCAAAAAGCAAGTATTCAATTTTGTGTTTAAAAATGTTTAAATGTTTAA 300
DB	241 GAAATGACCAAAAAGCAAGTATTCAATTTTGTGTTTAAAAATGTTTAAATGTTTAA 300
QY	301 GTGCTTAAACCAAGAATATATTTCTCGGTGATGTTTAAATGTTTGTGCAACATGATG 360
DB	301 GTGCTTAAACCAAGAATATATTTCTCGGTGATGTTTAAATGTTTGTGCAACATGATG 360
QY	361 GGAAGACCAAGCTGCGACTGCGATGCTAATTTGGAGGAAGGACTTTTGTCAAGCT 420
DB	361 GGAAGACCAAGCTGCGACTGCGATGCTAATTTGGAGGAAGGACTTTTGTCAAGCT 420
QY	421 CAAGGGAATGTTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGGTTGGTAAACAG 480
DB	421 CAAGGGAATGTTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGGTTGGTAAACAG 480
QY	481 TGTAATGTGCAACTAACACGAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT 540
DB	481 TGTAATGTGCAACTAACACGAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT 540
QY	541 GAGTGGGTACTCTCTACTTATAAGCATTAAGCAAAACCAAAAGAGCTATACCAAGTGT 600
DB	541 GAGTGGGTACTCTCTACTTATAAGCATTAAGCAAAACCAAAAGAGCTATACCAAGTGT 600
QY	601 GTTCTTTTGGAAACATGATTGCTTACTATTTTAACTTAAAGAAATAAGCAGCTAGT 660
DB	601 GTTCTTTTGGAAACATGATTGCTTACTATTTTAACTTAAAGAAATAAGCAGCTAGT 660
QY	661 CCACCAAGGAGCGAGCTATTTCTTAGCAGTGACTCTGCTGGGAAAACTAACTTTTAA 720
DB	661 CCACCAAGGAGCGAGCTATTTCTTAGCAGTGACTCTGCTGGGAAAACTAACTTTTAA 720
QY	721 AAAGAAGGCGAGCGCCATCTAGTGAGCAAACTATACCTGATGACATGCGGCGAGAAACG 780
DB	721 AAAGAAGGCGAGCGCCATCTAGTGAGCAAACTATACCTGATGACATGCGGCGAGAAACG 780
QY	781 GTTGAACCAACAGTAACCACTGCGCAGGAAACTAAGCGGCGAGAAATTCAACTTAAAA 840
DB	781 GTTGAACCAACAGTAACCACTGCGCAGGAAACTAAGCGGCGAGAAATTCAACTTAAAA 840
QY	841 GAAGTTTCTATTAAAACTACACTTAAAGAGCTGCTGCATAAAAAGAGTAACCTCACAGAG 900
DB	841 GAAGTTTCTATTAAAACTACACTTAAAGAGCTGCTGCATAAAAAGAGTAACCTCACAGAG 900
QY	901 GACTGGATGATGATGAGCAGCAGACAGTTACATTGAAATGATGCTCAACAGGTGGAGAA 960
DB	901 GACTGGATGATGATGAGCAGCAGACAGTTACATTGAAATGATGCTCAACAGGTGGAGAA 960
QY	961 AACTGCTGAAAAATACGCTAGAGATTGTACATTAACCTTAGCCAGAAACCAAAACAGCA 1020
DB	961 AACTGCTGAAAAATACGCTAGAGATTGTGTACATTAACCTTAGCCAGAAACCAAAACAGCA 1020
QY	1021 TTTGACTTAATTTTAAAAAGCTGAAAAACCAAGCAAACTAACCAACTTTTCACTGCCTGAC 1080
DB	1021 TTTGACTTAATTTTAAAAAGCTGAAAAACCAAGCAAACTAACCAACTTTTCACTGCCTGAC 1080
QY	1081 ACAAGAACCTGCAAGATTTTGTCTTTTCTGCTGGAACTATGTTAAAGTTTGCATGCT 1140
DB	1081 ACAAGAACCTGCAAGATTTTGTCTTTTCTGCTGGAACTATGTTAAAGTTTGCATGCT 1140
QY	1141 ATTTGCTGTGTTTTTAAACAGAGGAGGAAAAAGAAATGCTGTTTTTATTTTATGAGCA 1200
DB	1141 ATTTGCTGTGTTTTTAAACAGAGGAGGAAAAAGAAATGCTGTTTTTATTTTATGAGCA 1200
QY	1201 GCCAGACAGGCAAACTTATTATTGACAGCCATAGCAAGCAGTTGCAATGTTGGT 1260
DB	1201 GCCAGACAGGCAAACTTATTATTGACAGCCATAGCAAGCAGTTGCAATGTTGGT 1260
QY	1261 TGCTATAATGCAAGCAATGTAAACTTTTCCATTTTAAATGACGTGTACCAACCAAGAACT 1320
DB	1261 TGCTATAATGCAAGCAATGTAAACTTTTCCATTTTAAATGACGTGTACCAACCAAGAACT 1320
QY	1321 TGGGTAGAGAAAGCTGGTAACTTTTGGACAGCAAGTAAACAGTTTAAAGCAATTTGCTCT 1380
DB	1321 TGGGTAGAGAAAGCTGGTAACTTTTGGACAGCAAGTAAACAGTTTAAAGCAATTTGCTCT 1380
QY	1381 GGTCAAACTATTTCGCATTGATCAAAAGGAAAGGAGCAAGATTTGAACCAACCA 1440
DB	1381 GGTCAAACTATTTCGCATTGATCAAAAGGAAAGGAGCAAGATTTGAACCAACCA 1440
QY	1441 GTCATCATGACCAAAATGAGAACTATTACAGTGGTCAGAAATAGGCTCGAAGAAAGACCA 1500
DB	1441 GTCATCATGACCAAAATGAGAACTATTACAGTGGTCAGAAATAGGCTCGAAGAAAGACCA 1500
QY	1501 GAACACACTCAACCAATGAGAGACAGATGCTTAACTTAACTTAACTTAACTTAACTTAA 1560
DB	1501 GAACACACTCAACCAATGAGAGACAGATGCTTAACTTAACTTAACTTAACTTAACTTAA 1560
QY	1561 GGTGACCTTTGGTTGGTTCACAAAATGAATGCCCCATGATTTGCTGCTTGGTGGTAAAG 1620
DB	1561 GGTGACCTTTGGTTGGTTCACAAAATGAATGCCCCATGATTTGCTGCTTGGTGGTAAAG 1620

CDS	2354..2398	/gene="vp"	
		/note="ORF2; putative"	
		/codon_start=1	
		/product="unknown protein"	
		/protein_id="AA67113.1"	
		/db_xref="GI:825484"	
exon	2399..>4557	/translation="MFNYLFRPEITWF"	
		/gene="VPI"	
		/number=2	
CDS	2794..4557	/gene="VPI"	
		/note="VP2"	
		/codon_start=1	
		/protein_id="AA67114.1"	
		/db_xref="GI:332296"	
		/translation="MSDQTSQDPSGNVHSAERVAADGPGSGSGSGGGVGVST	
		GSYDQTHYFLDGGWVEITATLRLVHLMPSKSENYCRIRVNTDTSVKGNMAKDD	
		AHQEIWTPSWLDANAWQVLDQWQYICNTMSQNLVSLDQIEINVLKTVTEQDL	
		GGQAIKIYNNDLTACMMVAVDSSNNILPYTPAAMSMTFLGPKWPTIASPYRYFVD	
		RDLSTVENOEGTVEHVMGTGPKMNSQFFTIENQIOTILLRTGDFATGTVYFDRTNS	
		VKLTHWQTRQILOGPPLSTFPEADTDAGTLTAQGRSHGCTTQGMVNWVSEARTRPA	
		QVGFPCPHNDPEASRAGPFAAPKVPADITQGVDEKANGSVRYSGKHGHNWASHGPA	
		PERYTDWETSFSGRDTKGFIOASPLVPPPLNGILITNANPIGTQNDLHFSNVFNSY	
		GPLTAFSHPSPVYVQQQWIDKELDLHKPLRHITAPFVCKRNAPGQVQLVRLGNLTDQ	
Query Match	99.9%; Score 2017.4; DB 14; Length 5149;		
Best Local Similarity	100.0%; Pred. No. 0;		
Matches 2018; Conservative	0; Mismatches 1; Indels 0; Gaps 0;		
Qy	1	ATGCTGCGAAATGCTTACTCTGATGAAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA	60
Db	261	ATGGCTGAAATGCTTACTCTGATGAAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA	320
Qy	61	AGTAACGAGAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGNAAA	120
Db	321	AGTAACGAGAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGNAAA	380
Qy	121	GATATCGATGGAAATAGTTACAAAAAGAGCTGCAGAGGACGAGCTGAAATCTTTACAA	180
Db	381	GATATCGATGGAAATAGTTACAAAAAGAGCTGCAGAGGACGAGCTGAAATCTTTACAA	440
Qy	181	CGAGGACGGAATCTACTTGGGACCAAGCAGGACATGGAATGGGAAACCAACAGTGGAT	240
Db	441	CGAGGACGGAATCTACTTGGGACCAAGCAGGACATGGAATGGGAAACCAACAGTGGAT	500
Qy	241	GAAATGACCAAAAAGCAAGTATTCATTTTGGTTTGGTTTAAAAATGTTTATTGGAA	300
Db	501	GAAATGACCAAAAAGCAAGTATTCATTTTGGTTTGGTTTAAAAATGTTTATTGGAA	560
Qy	301	GTGCTTAACCAAGAATATATTTCTGGTGATCTTAATTTGGTTGGCAACATGAATGG	360
Db	561	GTGCTTAACCAAGAATATATTTCTGGTGATCTTAATTTGGTTGGCAACATGAATGG	620
Qy	361	GGAAAGACCAAGCTGGCAGTGCATGTACTAATTTGGAGGAAGGACTTTAGTCAAGCT	420
Db	621	GGAAAGACCAAGCTGGCAGTGCATGTACTAATTTGGAGGAAGGACTTTAGTCAAGCT	680
Qy	421	CAAGGGAATTTGGGAGAGCAACTAAATGTTTACTTGGAGCAGATGGTTGGTAAACAGCC	480
Db	681	CAAGGGAATTTGGGAGAGCAACTAAATGTTTACTTGGAGCAGATGGTTGGTAAACAGCC	740
Qy	481	TGTAATGTGCACTAAACACAGCTGAAGAATTAACCTAAGAGAAATAGCAGAGACAAT	540
Db	741	TGTAATGTGCACTAAACACAGCTGAAGAATTAACCTAAGAGAAATAGCAGAGACAAT	800
Qy	541	GAGTGGGTACTCTACTTACTTATAAGCATAAGCAAAACCAAAAGACTATACCAAGTGT	600
Db	801	GAGTGGGTACTCTACTTACTTATAAGCATAAGCAAAACCAAAAGACTATACCAAGTGT	860
Qy	601	GTTCCTTTTGGAAACATGATTTGCTTACTATTTTAACTAAAAAGAAAAATTAAGCACTAGT	660

Db	861	GTTCCTTTTGGAAACATGATTGCTTACTATTTTAACTAAAAAGAAAAATAAGCACTAGT	920
Qy	661	CCACCAAGAGACGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAAATACTAATTTTAA	720
Db	921	CCACCAAGAGACGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAAATACTAATTTTAA	980
Qy	721	AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACCTGATGACATCGCGCCAGAAACG	780
Db	981	AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACCTGATGACATCGCGCCAGAAACG	1040
Qy	781	GTTGAAACACACAGTAACCACTGCGCAGGAACTAAGCGCGGAGAAATTCAAAATAAAAAA	840
Db	1041	GTTGAAACACACAGTAACCACTGCGCAGGAACTAAGCGCGGAGAAATTCAAAATAAAAAA	1100
Qy	841	GAAAGTTTCTAATTAATACTACATTAAGAGCTGGTGATAAAAAGAGTAACCTCACAGAG	900
Db	1101	GAAAGTTTCTAATTAATACTACATTAAGAGCTGGTGATAAAAAGAGTAACCTCACAGAG	1160
Qy	901	GACTGATGATGATGAGCAGCAGACAGTTACATTTGAATGATGCTCAACAGCTGGAGAA	960
Db	1161	GACTGATGATGATGAGCAGCAGACAGTTACATTTGAATGATGCTCAACAGCTGGAGAA	1220
Qy	961	AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAAACAAAAACAGCA	1020
Db	1221	AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAAACAAAAACAGCA	1280
Qy	1021	TTTGACTTAATTTTAGAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCTGAC	1080
Db	1281	TTTGACTTAATTTTAGAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCTGAC	1340
Qy	1081	ACAAGAACCTGAGAAATTTTGGCTTTTTCATGGCTGGAACCTATGTTAAAGTTTGGCATGCT	1140
Db	1341	ACAAGAACCTGAGAAATTTTGGCTTTTTCATGGCTGGAACCTATGTTAAAGTTTGGCATGCT	1400
Qy	1141	ATTGCTGTGTTTTTAAACGACAGAGGCAAGAAAGAAATGCTGTTTTTATTTTATGAGACCA	1200
Db	1401	ATTGCTGTGTTTTTAAACGACAGAGGCAAGAAAGAAATGCTGTTTTTATTTTATGAGACCA	1460
Qy	1201	GCAGACAGGCAAAATCTATTATTGCAAGCCATAGCAAGCAGTGGCAATGTTGGT	1260
Db	1461	GCAGACAGGCAAAATCTATTATTGCAAGCCATAGCAAGCAGTGGCAATGTTGGT	1520
Qy	1261	TGCTATAATGCAGCCAAATGTAACTTTCCATTTTAAATGACTGTACCAACAGAACTTGATT	1320
Db	1521	TGCTATAATGCAGCCAAATGTAACTTTCCATTTTAAATGACTGTACCAACAGAACTTGATT	1580
Qy	1321	TGCGTAGAAGAACTGTTAACTTTTGACACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT	1380
Db	1581	TGCGTAGAAGAACTGTTAACTTTTGACACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT	1640
Qy	1381	GGTCAAACTATTTCGCATTGATCAAAAAGGAAAGGAGCAGCAAAAGAGATTGAACCAACCA	1440
Db	1641	GGTCAAACTATTTCGCATTGATCAAAAAGGAAAGGAGCAGCAAAAGAGATTGAACCAACCA	1700
Qy	1441	GTCAATCATGACCAAAATGAGAACTTAAGTGGTCAAGATAGGCTGCGAAGAAAGACCA	1500
Db	1701	GTCAATCATGACCAAAATGAGAACTTAAGTGGTCAAGATAGGCTGCGAAGAAAGACCA	1760
Qy	1501	GAACACACTCAACCAATCAGAGACAGATGCTTAACTTAACTTAACTTAACTTAACTTAACT	1560
Db	1761	GAACACACTCAACCAATCAGAGACAGATGCTTAACTTAACTTAACTTAACTTAACTTAACT	1820
Qy	1561	GGTGACTTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGTGCTGCTGGTAAAG	1620
Db	1821	GGTGACTTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGTGCTGCTGGTAAAG	1880
Qy	1621	AATGGTTTAACTTCACTACCTGCAAGCTACTGTCTTAAATGGGGCAAAAGTTCTCTGATGG	1680
Db	1881	AATGGTTTAACTTCACTACCTGCAAGCTACTGTCTTAAATGGGGCAAAAGTTCTCTGATGG	1940
Qy	1681	TCAGAAAACTGGCGGAGCAAGGTCGCAACTTCTTAAATTTTAACTTAACTTAACTTAACT	1740
Db	1941	TCAGAAAACTGGCGGAGCAAGGTCGCAACTTCTTAAATTTTAACTTAACTTAACTTAACT	2000

```
QY 1741 TCACCAITTCAGACCCGAAAGTAGCGCTCTCAGCCAGAGAACTATGCACTAACTCCACTT 1800
DB 2001 TCACCAITTCAGACCCGAAAGTAGCGCTCTCAGCCAGAGAACTATGCACTAACTCCACTT 2060
QY 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGGAGCACACCAAAATACCTCTGTTGGC 1860
DB 2061 GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGGAGCACACCAAAATACCTCTGTTGGC 2120
QY 1861 GGCACCTGAGAAACCCAGAACACTGGGGAGAGCTGGTCCAAAGCTGCGCAAGATGGTCAA 1920
DB 2121 GGCACCTGAGAAACCCAGAACACTGGGGAGAGCTGGTCCAAAGCTGCGCAAGATGGTCAA 2180
QY 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGATTTGAGAGCGTCTCGGTGGCGAACC 1980
DB 2181 CTGAGCCCAACTTGGTCAGAGATCGAGGAGATTTGAGAGCGTCTCGGTGGCGAACC 2240
QY 1981 TTGAAGAAAGACTTCAGGAGCCGCTGAATTTGACTAA 2019
DB 2241 TTGAAGAAAGACTTCAGGAGCCGCTGAATTTGACTAA 2279

RESULT 5
AX137739
LOCUS AX137739 2019 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 4 from Patent EP1077260.
ACCESSION AX137739
VERSION AX137739.1 GI:14273913
KEYWORDS
SOURCE
ORGANISM Mice minute virus
REFERENCE Mice minute virus
AUTHORS Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
Nuesch, J. and Rommelaere, J.
TITLE Parvovirus ns1 variants
JOURNAL Patent: EP 1077260-A 4 21-FEB-2001;
DEUTSCHES KREBSFORSCHUNGSZENTRUM STIFTUNG DES OEFFENTLICHEN RECHTS
(DE)

FEATURES
source location/Qualifiers
1..2019
/organism="Mice minute virus"
/mol_type="unassigned DNA"
/db_xref="taxon:10794"
/notes="Parvovirus NS1 variant"
1..2019
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC39991.1"
/db_xref="GI:14273914"
/translacion="MAGNAYSDEVLTGATNMLKEKSQEVFSFVKQNVOLNGKDIGW
NSYKKELOEDELKSLQGAETTWQSDSEDMEWETVDEMTKKQVFIIDSLVKKCLPEVL
NTKNIIPGDVNVFWFHEWKGDKGCHVLIGKDFSOAGKWRRLNVYWSRLVTA
CNVQLTPAERIKLREIADNEWVLLTYKHQTKDYTKCVLFQNMIAFYFLTKKIS
TSPRDGGYFLSSDSGKMTNFKEGEHLVSKLYTDDRPETVEITVTVAQETPKRRI
QTKKVAIKTLKELVHKRVTSPEDMMWQDSYIEMMAQPGGENLLKNTLEICTLTL
ARTKAFDLILEAKETSKLNFSLPDTTRCIRIIFAFHGNVYKVCNHAICCVLNRQGGK
NTVLFHPASTGKSI1AQAOAGVNGVCYNAANVPFPNDCTNKLIVVEBAQNGFQ
QVNOFKAIQSGOTIRIDQKGSQKLEPTPVINTNENITVVRIGCEERPEHTQPIRD
RMLNIHLTHPLRGDFGLVDKNEPMICAMLVKNGYQSTMASYCAKWKVQPDSENAE
PKVPTPINLGSARSPTTIPKSTNTALTDLEADLEPDEPSTENIPVAGTAE
TQNTGEAGSKACQDQGLSPSTWSEIEDLRACFGAEPLKDFSEPLND"

ORIGIN
Query Match 99.8%; Score 2015.8; DB 6; Length 2019;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGCTCGAAATGCTTACTCTGATGAAGTTTGGAGCAACCACTGGTTTAAAGGAAAA 60
DB 1 ATGCTCGAAATGCTTACTCTGATGAAGTTTGGAGCAACCACTGGTTTAAAGGAAAA 60
QY 61 AGTACAGGAGAGTCTCATTTGTTTTTAAAAATGCAATGTTCAACTGAATGAAAA 120
```

```
DB 61 AGTAAACGAGGAGTGTCTCATTTGTTTTTAAAAATGAAAAATGTTCAACTGAATGAAAA 120
QY 121 GATATCGGATGGAAATAGTTTACAAAAGAGCTCAGAGGAGAGCTGGAATCTTTTACAA 180
DB 121 GATATCGGATGGAAATAGTTTACAAAAGAGCTCAGAGGAGAGCTGGAATCTTTTACAA 180
QY 181 CGAGGAGCGGAAATCTACTTTGGGACCAAGCAGGACATGGAATGGGAAACCAACAGTGGAT 240
DB 181 CGAGGAGCGGAAATCTACTTTGGGACCAAGCAGGACATGGAATGGGAAACCAACAGTGGAT 240
QY 241 GAAATGACCAAAAGCAAGATTAATTTGTTGATTTCTTTGTTTAAAAATGTTTATTGAA 300
DB 241 GAAATGACCAAAAGCAAGATTAATTTGTTGATTTCTTTGTTTAAAAATGTTTATTGAA 300
QY 301 GTGCTTAAACAAAGAAATATATTTCTGTTGATTTAAATTTGTTTAAACATGAAATGG 360
DB 301 GTGCTTAAACAAAGAAATATATTTCTGTTGATTTAAATTTGTTTAAACATGAAATGG 360
QY 361 GGAAGAGACCAAGCTGGCCTGCACTGCAATGCTAATTTGGAGGAAAGGACTTTTCAAGCT 420
DB 361 GGAAGAGACCAAGCTGGCCTGCACTGCAATGCTAATTTGGAGGAAAGGACTTTTCAAGCT 420
QY 421 CAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTGAGCAGATGGTTGTTAAACAGCC 480
DB 421 CAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTGAGCAGATGGTTGTTAAACAGCC 480
QY 481 TGTAAATGTCAACTAACAACAGCTGAAAGAAATTTAAACTAAGGAAATAGCAGAGCAAT 540
DB 481 TGTAAATGTCAACTAACAACAGCTGAAAGAAATTTAAACTAAGGAAATAGCAGAGCAAT 540
QY 541 GAGTGGGTTTACTCTACTTCTTATAAGCATAAGCAACCAACCAAGGACTATACCAAGTGT 600
DB 541 GAGTGGGTTTACTCTACTTCTTATAAGCATAAGCAACCAACCAAGGACTATACCAAGTGT 600
QY 601 GTTCTTTTGGAAAACATGATTTGCTTACTTATTTTAACTTAAAGAAAATAGCACTAGT 660
DB 601 GTTCTTTTGGAAAACATGATTTGCTTACTTATTTTAACTTAAAGAAAATAGCACTAGT 660
QY 661 CCACCAAGAGAGCGAGGCTATTTCTTAGCAGTACTCTGGCTGGGAAACTTAACTTTTAA 720
DB 661 CCACCAAGAGAGCGAGGCTATTTCTTAGCAGTACTCTGGCTGGGAAACTTAACTTTTAA 720
QY 721 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATCGGCGCAGAAACG 780
DB 721 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATCGGCGCAGAAACG 780
QY 781 GTTGAACACACAGTAAACCACTGCGCAGGAAACTTAAAGCGCGGAGAAATCAAACTAAAAA 840
DB 781 GTTGAACACACAGTAAACCACTGCGCAGGAAACTTAAAGCGCGGAGAAATCAAACTAAAAA 840
QY 841 GAAATTTCTTATTTAAAACTACACTTAAAGAGCTGGTGTCATAAAAAGAGTAACTCACCAGAG 900
DB 841 GAAATTTCTTATTTAAAACTACACTTAAAGAGCTGGTGTCATAAAAAGAGTAACTCACCAGAG 900
QY 901 GACTGGATGATGATGAGCAGCAGACAGTTTACATTTGAAATGATGGCTCAACAGGTGGAGAA 960
DB 901 GACTGGATGATGATGAGCAGCAGACAGTTTACATTTGAAATGATGGCTCAACAGGTGGAGAA 960
QY 961 AACCTGCTGAAAAATACGCTAGAGATTTCTACACTAACTCTAGCCAGAACCAACAGCA 1020
DB 961 AACCTGCTGAAAAATACGCTAGAGATTTCTACACTAACTCTAGCCAGAACCAACAGCA 1020
QY 1021 TTTGACTTAAATTTAGAAAAAGCTGAAACACAGCAAACTAACTTTTCACTGCTGTAC 1080
DB 1021 TTTGACTTAAATTTAGAAAAAGCTGAAACACAGCAAACTAACTTTTCACTGCTGTAC 1080
QY 1081 ACAAGAACCTGAGAAATTTTCTTTTATGCTGGAACATGTTTAAAGTTTGCATGCT 1140
DB 1081 ACAAGAACCTGAGAAATTTTCTTTTATGCTGGAACATGTTTAAAGTTTGCATGCT 1140
QY 1141 ATTTGCTGTGTTTTTAAACAGACCAAGGAGCAAAAGAAATGCTGTTTATTTTCAATGACCA 1200
DB 1141 ATTTGCTGTGTTTTTAAACAGACCAAGGAGCAAAAGAAATGCTGTTTATTTTCAATGACCA 1200
```

Qy	1201	GCACGACAGGCAAACTTATTATGCAACAGCCATAGCAACAGCAGTTGGCAATGTTGGT	1260
Db	1201	GCACGACAGGCAAACTTATTATGCAACAGCCATAGCAACAGCAGTTGGCAATGTTGGT	1260
Qy	1261	TGCTATAATCAGCAGCAATGTAATCTTCCATTTAAGACTGTACCAACAGAACTTGATT	1320
Db	1261	TGCTATAATCAGCAGCAATGTAATCTTCCATTTAAGACTGTACCAACAGAACTTGATT	1320
Qy	1321	TGGGTAGAAGAGCTGGTAACTTTGGACAGCAAGTAACACAGTTTAAAGCCATTGCTCT	1380
Db	1321	TGGGTAGAAGAGCTGGTAACTTTGGACAGCAAGTAACACAGTTTAAAGCCATTGCTCT	1380
Qy	1381	GGTCAAACTATTGCGATTGATTCAAAGAGAAAGGAGGAGCAACAGATTGAACCAACCA	1440
Db	1381	GGTCAAACTATTGCGATTGATTCAAAGAGAAAGGAGGAGCAACAGATTGAACCAACCA	1440
Qy	1441	GTCAATCATGACCAACAAATGAGACATTACAGTGTGTAGATAGGCTGCGAAGAAAGACCA	1500
Db	1441	GTCAATCATGACCAACAAATGAGACATTACAGTGTGTAGATAGGCTGCGAAGAAAGACCA	1500
Qy	1501	GAACACACTCAACCAATCAGAGACAGAACTTAAACATTCATCTAACACATACCTTGCCCT	1560
Db	1501	GAACACACTCAACCAATCAGAGACAGAACTTAAACATTCATCTAACACATACCTTGCCCT	1560
Qy	1561	GGTGACTTTGGTTGGTTGCAAAATGAATGGCCCAATGATTGCTGGTTGGTAAAG	1620
Db	1561	GGTGACTTTGGTTGGTTGCAAAATGAATGGCCCAATGATTGCTGGTTGGTAAAG	1620
Qy	1621	AATGGTTACCAATCTACCATGGCAAGCTACTGTCTTAATGGGCAAGTTCTCTGATTGG	1680
Db	1621	AATGGTTACCAATCTACCATGGCAAGCTACTGTCTTAATGGGCAAGTTCTCTGATTGG	1680
Qy	1681	TCAGAAAACCTGGCGGAGCCAAAGTGCCAACTCTATAAATTAATAGTTCGGCACGC	1740
Db	1681	TCAGAAAACCTGGCGGAGCCAAAGTGCCAACTCTATAAATTAATAGTTCGGCACGC	1740
Qy	1741	TCACCAATTCACGACACCGAAAAGTACGCTCTCAGCCAGAACTATGCACTAACCTCACTT	1800
Db	1741	TCACCAATTCACGACACCGAAAAGTACGCTCTCAGCCAGAACTATGCACTAACCTCACTT	1800
Qy	1801	GCATCGGATCTCAGGACCTGGCTTTAGAGCCTTGGAGCACCAATATCTCTGTTGG	1860
Db	1801	GCATCGGATCTCAGGACCTGGCTTTAGAGCCTTGGAGCACCAATATCTCTGTTGG	1860
Qy	1861	GGCACTGCAGAAACCCAGAACTACCTGGGAAAGCTGGTTCCAAAGCTGCCAAGATGGTCAA	1920
Db	1861	GGCACTGCAGAAACCCAGAACTACCTGGGAAAGCTGGTTCCAAAGCTGCCAAGATGGTCAA	1920
Qy	1921	CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTGAGAGCGTCTTCGGTGGGAAACCG	1980
Db	1921	CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTGAGAGCGTCTTCGGTGGGAAACCG	1980
Qy	1981	TTGAAGAAAGACTTCAGCGAGCGCTGAACCTGGACTAA	2019
Db	1981	TTGAAGAAAGACTTCAGCGAGCGCTGAACCTGGACTAA	2019

RESULT 6
AX137743
LOCUS
DEFINITION
Sequence 8 from Patent EP1077260.
ACCESSION
AX137743
VERSION
AX137743.1
KEYWORDS
GI:14273917
SOURCE
Mice minute virus
ORGANISM
Mice minute virus
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
1
REFERENCE
Nueesch, J. and Rommelaere, J.
Parvovirus ns1 variants
TITLE
Patent: EP 1077260-A 8 21-FEB-2001;
JOURNAL
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts

(DB)

FEATURES
source

Location/Qualifiers
1. .2019
/organism="Mice minute virus"
/mol_type="unassigned DNA"
/db_xref="taxon:10794"
/note="Parvovirus NS1 variant"
1. .2019
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC39993.1"
/db_xref="GI:14273918"
/translation="MAGNAYSDEVLCATNWLKESNQEVESFVKHENVOLNKGDKTGW
NSTKELQDELKSLQAGAEITDQSDMEWETTVDEMTEKQVIFDLSLVKCLFVFL
NTKNTIPFGVNMVFWHEGWDQGHCHVLIGKDFSOAQKWRRLNVLNYSWMLVTA
CNVQLTPAERIKLREIAEDNEWLLTYKHQTKDYTKCVLFNNIAYYFLTKKIS
TSPRPDGGYRIKLSRDEGGERHLVSKLYTDDMRPETVTTVTAQETKGRIR
QTKKEVSIKNTLKELVHKTVPEDMMQPDYSIEMMAQPGGENLLKNTLEICTLTL
ARTKTAFDLILLEKETSKLNTSPSLPTRACRIPAFHGNVYKVKCHALCCVLRQGGK
NTVLFHGPASTGSIITQAIAQAVGNVGTNAANVFPNDKNTKMLVWEEAGNFGQ
QVNFKAICSGQIRIDOKGKSKQIEPTFVINTNENITVVRIGCEERPEHTQPIRD
RMLNIHLTHLTFGLVDKNWPMICAWLVKNGYQSTWASVCAKMKVDPDSENAE
PKVPTPINLLGSARSPTTPKSTPLSONVALTPLASDLDELALPEWSTNTPVAGTAE
TQNTGBAGSKACODGQLSPTWSEIEBDELRACFAGELPKKDFSEPLND"

ORIGIN

Query Match	99.8%;	Score	2015.8;	DB	6;	Length	2019;
Best Local Similarity	99.9%;	Pred. No.	0;				
Matches	2017;	Conservative	0;	Mismatches	2;	Indels	0;
Gaps	0;						
QY	1	ATGGCTGGAAAATGCTTACTCTGATGAAATGTTTGGGAGCAACCAACTGGTTTAAAGGAAAAA	60				
Db	1	ATGGCTGGAAAATGCTTACTCTGATGAAATGTTTGGGAGCAACCAACTGGTTTAAAGGAAAAA	60				
QY	61	AGTAACCCAGGAAAGTGTCTCATTTGTTTAAAAATGAAAAATGTTTCAACTGAATGGAAAAA	120				
Db	61	AGTAACCCAGGAAAGTGTCTCATTTGTTTAAAAATGAAAAATGTTTCAACTGAATGGAAAAA	120				
QY	121	GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTTACAA	180				
Db	121	GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTTACAA	180				
QY	181	CGAGGAGCGGAAACTACTTTGGGACCAAGCGAGGACATGGAAATGGGAAACCAACAGTGGAT	240				
Db	181	CGAGGAGCGGAAACTACTTTGGGACCAAGCGAGGACATGGAAATGGGAAACCAACAGTGGAT	240				
QY	241	GAAATGACCAAAAGCAAGTATTCTTTTGTGATCTTTTGGTTAAAAAATGTTTTATTGTAA	300				
Db	241	GAAATGACCAAAAGCAAGTATTCTTTTGTGATCTTTTGGTTAAAAAATGTTTTATTGTAA	300				
QY	301	GTGCTTTAACACAAAGAAATATATTTCTGCTGATGTTAAATTTGGTTTGTGCACATGAATGG	360				
Db	301	GTGCTTTAACACAAAGAAATATATTTCTGCTGATGTTAAATTTGGTTTGTGCACATGAATGG	360				
QY	361	GGAAAAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGGACTTTTACTCAAGCT	420				
Db	361	GGAAAAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGGACTTTTACTCAAGCT	420				
QY	421	CAAGGGAAATGTTGGAGAAAGGCAACTAAATGTTTACTTGGAGCAGATGTTTGGTTAAACAGCC	480				
Db	421	CAAGGGAAATGTTGGAGAAAGGCAACTAAATGTTTACTTGGAGCAGATGTTTGGTTAAACAGCC	480				
QY	481	TGTAATGTGCAACTAACACAGCTGAAAGAAATTTAACTAAGAGAAATAGCAGAGACAAT	540				
Db	481	TGTAATGTGCAACTAACACAGCTGAAAGAAATTTAACTAAGAGAAATAGCAGAGACAAT	540				
QY	541	GAGTGGGTTTACTTCTACTTATAGCATAGCAAAACCAAAAAAGACTATATCAAGTGT	600				
Db	541	GAGTGGGTTTACTTCTACTTATAGCATAGCAAAACCAAAAAAGACTATATCAAGTGT	600				
QY	601	GTTCTTTTTGGAAACATGATGCTTACTATTTTTTAACTAAAAAGAAAAATAAGCACTAGT	660				
Db	601	GTTCTTTTTGGAAACATGATGCTTACTATTTTTTAACTAAAAAGAAAAATAAGCACTAGT	660				

121 GATATCGGATGGAATAGTTTACAAAAAGAGCTCGAGGAGCAAGCTGTAATCTTTACAA 180
121 GATATCGGATGGAATAGTTTACAAAAAGAGCTCGAGGAGCAAGCTGTAATCTTTACAA 180
181 CGAGGAGCGGAATCTACTTGGGACCAAGCAGGAGCATGAATGGGAAACCCACAGTGGAT 240
181 CGAGGAGCGGAATCTACTTGGGACCAAGCAGGAGCATGAATGGGAAACCCACAGTGGAT 240
241 GAAATGACAAAAAGCAAGATTTTCAATTTTGTATTTTGTGTTAAAAAATGTTTATTGAA 300
241 GAAATGACAAAAAGCAAGATTTTCAATTTTGTATTTTGTGTTAAAAAATGTTTATTGAA 300
301 GTCTTTAAACAACAAGAAATATTTTCTGTTGATTTGTTAAATTTGTTGTGCAACATGAATGG 360
301 GTCTTTAAACAACAAGAAATATTTTCTGTTGATTTTGTGTTAAATTTGTTGTGCAACATGAATGG 360
361 GGAAGAAGCAAGCTGGCACTGCCATGTACTTAATTTGGAGGAAGCACTTTAGTCAAGCT 420
361 GGAAGAAGCAAGCTGGCACTGCCATGTACTTAATTTGGAGGAAGCACTTTAGTCAAGCT 420
421 CAAGGGAAATGGTGGAGAAAGCACTAAATGTTTACTTGGAGCAGATGTTTGGTAAACAGCC 480
421 CAAGGGAAATGGTGGAGAAAGCACTAAATGTTTACTTGGAGCAGATGTTTGGTAAACAGCC 480
481 TGTAAATGTCAAACTAAACAAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGAACAAAT 540
481 TGTAAATGTCAAACTAAACAAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGAACAAAT 540
541 GAGTGGGTTACTCTTAATTAAGCATAAGCAAAACCAAAAAAGACTATATACCAAGTGT 600
541 GAGTGGGTTACTCTTAATTAAGCATAAGCAAAACCAAAAAAGACTATATACCAAGTGT 600
601 GTTCTTTTGGAAACATGATTTGCTTACTATTTTAACTAAAGAAATTAAGCACTAGT 660
601 GTTCTTTTGGAAACATGATTTGCTTACTATTTTAACTAAAGAAATTAAGCACTAGT 660
661 CCACCAAGAGACGGAGGCTATTTCTTAGCAGTACTCTGCTGGAATAACTTAACTTTTAA 720
661 CCACCAAGAGACGGAGGCTATTTCTTAGCAGTACTCTGCTGGAATAACTTAACTTTTAA 720
721 AAGAAGGAGCGGAGCCATCTAGTGTAGCAAACTATATACCTGATGATGATGCGGCGCAAGACG 780
721 AAGAAGGAGCGGAGCCATCTAGTGTAGCAAACTATATACCTGATGATGATGCGGCGCAAGACG 780
781 GTTGAACACACAGTAAACCACTGCGCAGGAACTAAGCGCGCAGAAATTCAACTTAAATAA 840
781 GTTGAACACACAGTAAACCACTGCGCAGGAACTAAGCGCGCAGAAATTCAACTTAAATAA 840
841 GAAGTTTCTATTTAACTTAACTTAAAGAGCTGGTGCATATAAGAGTAACTCACCAGAG 900
841 GAAGTTTCTATTTAACTTAACTTAAAGAGCTGGTGCATATAAGAGTAACTCACCAGAG 900
901 GACTGGATGATGATGACGAGCAGTATACCTGAAATGATGCTCAACAGGTTGGAGAA 960
901 GACTGGATGATGATGACGAGCAGTATACCTGAAATGATGCTCAACAGGTTGGAGAA 960
961 AACCTGCTGAAAAATAGCTAGAGATTTGTACACTTACTAGCTAGCAACCAAAACAGCA 1020
961 AACCTGCTGAAAAATAGCTAGAGATTTGTACACTTACTAGCTAGCAACCAAAACAGCA 1020
1021 TTTGACTTAAATTTTAAAAAGCTGAAACACGCAAACTAACAACCTTTTCACTGCTGAC 1080
1021 TTTGACTTAAATTTTAAAAAGCTGAAACACGCAAACTAACAACCTTTTCACTGCTGAC 1080
1081 ACAAGAACCTGCAGAAATTTTGTCTTTTCTGCTGGAATATGTTAAAGTTTGCATGCT 1140
1081 ACAAGAACCTGCAGAAATTTTGTCTTTTCTGCTGGAATATGTTAAAGTTTGCATGCT 1140
1141 ATTTGCTGTGTTTAAACAGACAGAGGAGGCAAAAGAAATGCTGTTTATTTATGAGACCA 1200
1141 ATTTGCTGTGTTTAAACAGACAGAGGAGGCAAAAGAAATGCTGTTTATTTATGAGACCA 1200

QY 1201 GCCAGCAGGCGAATCTATTATTCACAAGCCATAGACACAAGCAGTTGCGAATGTTGCT 1260
Db 1201 GCCAGCAGGCGAATCTATTATTCACAAGCCATAGACACAAGCAGTTGCGAATGTTGCT 1260
QY 1261 TGTATAATGTCAGCAATGTAACCTTTCCATTTTAAATGACTGTACCAACAGAACTTGATT 1320
Db 1261 TGTATAATGTCAGCAATGTAACCTTTCCATTTTAAATGACTGTACCAACAGAACTTGATT 1320
QY 1321 TGGGTAGAAAGAGCTGGTAACTTTTGGACAGCAAGTAACAGTTTAAAGCCATTGCTCT 1380
Db 1321 TGGGTAGAAAGAGCTGGTAACTTTTGGACAGCAAGTAACAGTTTAAAGCCATTGCTCT 1380
QY 1381 GGTCAAACTATTTCGCATTCATCAAAAGGAAAGGAGGAGCAAGATTAACCAACCA 1440
Db 1381 GGTCAAACTATTTCGCATTCATCAAAAGGAAAGGAGGAGCAAGATTAACCAACCA 1440
QY 1441 GTCATCATGACCAACCAATGAGAACATTACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA 1500
Db 1441 GTCATCATGACCAACCAATGAGAACATTACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA 1500
QY 1501 GAACACACTCAACCAATGAGAACATTACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA 1560
Db 1501 GAACACACTCAACCAATGAGAACATTACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA 1560
QY 1561 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTTGTGCTTGGTTGTTAAAG 1620
Db 1561 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTTGTGCTTGGTTGTTAAAG 1620
QY 1621 AATGGTTTACCAATCTACCATGCAAGCTACTGTCTTAAATGGGGCAAAAGTTCTCGATTGG 1680
Db 1621 AATGGTTTACCAATCTACCATGCAAGCTACTGTCTTAAATGGGGCAAAAGTTCTCGATTGG 1680
QY 1681 TCAGAAAACTGGCGGAGGCAAAAGTGGCCAACTCTTATAAATTTACTAGTTTCGCAAGC 1740
Db 1681 TCAGAAAACTGGCGGAGGCAAAAGTGGCCAACTCTTATAAATTTACTAGTTTCGCAAGC 1740
QY 1741 TCACCAATTCAGACACCGAAAGTAGCTCTCAGCAGAACATATGCACACTTCCACTT 1800
Db 1741 TCACCAATTCAGACACCGAAAGTAGCTCTCAGCAGAACATATGCACACTTCCACTT 1800
QY 1801 GCATCGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGACACACCAAACTACTCTGTTGCG 1860
Db 1801 GCATCGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGACACACCAAACTACTCTGTTGCG 1860
QY 1861 GGCACTGCAGAAACCCAGAACACTGGGGAGCTGGTTCCAAAGCCCTGCCAAGATGTTCAA 1920
Db 1861 GGCACTGCAGAAACCCAGAACACTGGGGAGCTGGTTCCAAAGCCCTGCCAAGATGTTCAA 1920
QY 1921 CTGAGCCCAACTTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGTTGCGGACCG 1980
Db 1921 CTGAGCCCAACTTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGTTGCGGACCG 1980
QY 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGAGACTAA 2019
Db 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGAGACTAA 2019

RESULT 8

MMVIGG 5085 bp ss-DNA linear VRL 05-JUL-1995
LOCUS Minute virus of mice (MMV(i)), a lymphotropic variant of MMV,
DEFINITION complete genome.
ACCESSION M12032
VERSION M12032.1 GI:332289
KEYWORDS alternative splicing; capsid protein; complete genome;
nonstructural protein.
SOURCE Mouse minute virus
ORGANISM Mouse minute virus
REFERENCE 1 (bases 1 to 5085)
AUTHORS Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
Astell,C.R., Gardiner,E.M. and Tattersall,P.
TITLE DNA sequence of the lymphotropic variant of minute virus of mice,
MMV(i), and comparison with the DNA sequence of the fibrotropic

ORIGIN	1085 bp upstream of EcoRI site.	
Query Match	94.4%; Score 1905.4; DB 14; Length 5085;	
Best Local Similarity	96.5%; Pred. No. 0;	
Matches 1948;	Conservative 0; Mismatches 71; Indels 0; Gaps 0;	
Qy	1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAAA 60	
Db	262 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAAACAACCACTGGTTAAAGGAAAAA 321	
Qy	61 AGTAACACGAGAGTGTCTCATTTGTTTTTAAATAAGTAATGTTCAACTGAATGGAATA 120	
Db	322 AGTAACACGAGAGTGTCTCATTTGTTTTTAAATAAGTAATGTTCAACTGAATGGAATA 381	
Qy	121 GATATCGGATGGAAATAGTTTACAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTACAA 180	
Db	382 GATATCGGATGGAAATAGTTTACAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTACAA 441	
Qy	181 CGAGGAGCGAAACTACTTGGGACCAAGGAGGACATGGAATGGGAAACCAAGTGGAT 240	
Db	442 CGAGGAGCGAAACTACTTGGGACCAAGGAGGACATGGAATGGGAACTACAGTGGAT 501	
Qy	241 GAAATGACCAAAAGCAAGTATTTCATTTTGGTTTGGTTTAAATAAGTGTATTATTGAA 300	
Db	502 GAAATGACCAAAAGCAAGTATTTCATTTTGGTTTGGTTTAAATAAGTGTATTATTGAA 561	
Qy	301 GTGCTTAAACAAAGAAATATATTTCTCGTGTGATGTTAAATGGTTTGTGCAACATGAATGG 360	
Db	562 GTGCTTAGCACAAAATATAGCTCTGCTGATGTTACTTTGTTGTGCAGCATGAATGG 621	
Qy	361 GGAAGAACCAAGCTGGCACTGCGCATGTAATTTGGAGGAAAGGACTTTAGTCAAGCT 420	
Db	622 GGGAAAGAACCAAGCTGGCACTGCGCATGTAATTTGGAGGCAAGGACTTTAGTCAAGCT 681	
Qy	421 CAAGGGAATGTTGGAGGCAAGCACTAAATGTTTACTTGGAGCAGATGGTTGGTAAACAGCC 480	
Db	682 CAAGGGAATGTTGGAGGCAAGCACTAAATGTTTACTTGGAGCAGATGGTTGGTAAACAGCC 741	
Qy	481 TGTAAATGTCAACTAAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGACAAT 540	
Db	742 TGTAAATGTGAGCTAAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGACAGT 801	
Qy	541 GAGTGGTTACTCTTACTTATAGCATAAAGCAAAACCAAAAAAGACTATATACCAAGTGT 600	
Db	802 GAGTGGTTACTCTTACTTATAGCATAAAGCAAAACCAAAAAAGACTATATACCAAGTGT 861	
Qy	601 GTTCTTTTGGAAATGATGTTCTTACTATTTTAAAGGAAATTAAGCACTAGT 660	
Db	862 GTTCTTTTGGAAATGATGTTCTTACTATTTTAAAGGAAATTAAGCACTAGT 921	
Qy	661 CCACCAAGAGACGGAGGCTATTTTCTAGCAGTGAATCTGGCTGGAAAACTAACTTTTAA 720	
Db	922 CCACCAAGAGACGGAGGCTATTTTCTAAGCAGTGAATCTGGCTGGAAAACTAACTTTTAA 981	
Qy	721 AAAGAGCGGAGCGCCATCTAGTGAGCAAACTATATACATGATGACATGCGGCGCAGAAACG 780	
Db	982 AAAGAGCGGAAACGCCATCTAGTGAGCAAACTATATACATGATGACATGCGGCGCAGAAACG 1041	
Qy	781 GTTGAACACCAAGTAACCACTGCGCAGGAACTAAGCGCGCGAGAAATTCAACTTAAATA 840	
Db	1042 GTTGAACACCAAGTAACCACTGCGCAGGAACTAAGCGCGCGAGAAATTCAACTTAAATA 1101	
Qy	841 GAAGTTTCTATTAATACTACATTTAAAGAGCTGGTGATTAAGAGAGTAACCTCACAGAG 900	
Db	1102 GAGTTTCTATTAATACTACATTTAAAGAGCTGGTGATTAAGAGAGTAACCTCACAGAA 1161	
Qy	901 GACTGGATGATGTCAGGCGACAGATTACATTGAAATGATGGCTCAACGAGTGGAGAA 960	
Db	1162 GACTGGATGATGTCAGGCGACAGATTACATTGAAATGATGGCTCAACGAGTGGAGAA 1221	
Qy	961 AACCTGCTGAAAAATAGCTAGAGATTGTACACTAATCTAGCCAGAAACCAAAACAGCA 1020	
Db	1222 AACCTGCTGAAAAATAGCTAGAGATTGTACGTAATCTAGCCAGAAACCAAAACAGCA 1281	

RESULT 9
MOU34253
LOCUS

MOU34253 4764 bp DNA linear VRL 21-AUG-1996

Qy	1021 TTTGACTTAATTTTAAAAAAGCTGAAACAGCAAACTAACCAACTTTTCACTGCTGAC 1080	
Db	1282 TTTGACTTGAATTTTAAAAAAGCTGAAACAGCAAACTAACCAACTTTTCACTGCTGAC 1341	
Qy	1081 ACAGAACTCTGCAAGATTTTTCCTTTTCATGCTGCAATATGTTTAAAGTTTCCCATGCT 1140	
Db	1342 ACAAGAACTCTGCAAGATTTTTCCTTTTCATGCTGCAATATGTTTAAAGTTTCCCATGCT 1401	
Qy	1141 ATTTGCTGTGTTTTTAAACAGACAAAGGAGCAAAAGAAATGCTGTTTATTTTCAATGACCA 1200	
Db	1402 ATTTGCTGTGTTTTTAAACAGACAAAGGAGCAAAAGAAATGCTGTTTATTTTCAATGACCA 1461	
Qy	1201 GCAGACACAGGCAAAATCTATTTATTTGCAAGCCATAGCAACAGCTTTGGCAATGTTGGT 1260	
Db	1462 GCAGTACAGGCAAAATCTATTTATTTGCAAGCCATAGCAACAGCTTTGGTAAATGTTGGT 1521	
Qy	1261 TGCCTAATATGCAAGCAATGTAAACTTTCCATTTTAAATGACTGTACCAACAGAACTTGAAT 1320	
Db	1522 TGCCTAATATGCAAGCAATGTAAACTTTCCATTTTAAATGACTGTACCAACAGAACTTGAAT 1381	
Qy	1321 TGGGTAGAAAGAGCTGGTAACTTTTGGACAGCAAGTAAACCCAGTTTAAAGCCATTGCTCT 1380	
Db	1582 TGGGTAGAAAGAGCTGGTAACTTTTGGACAGCAAGTAAACCCAGTTTAAAGCCATTGCTCT 1641	
Qy	1381 GGTCAAACTATTTCGCATTTGATCAAAAAGGAAAGGAGCAACAGATTTGAACCAACCA 1440	
Db	1642 GGTCAAACTATTTCGCATTTGATCAAAAAGGAAAGGAGCAACAGATTTGAACCAACCA 1701	
Qy	1441 GTCATCATGACCAAAATGAGAACATTCAGTGGTGCAGAAATAGGCTGCGAAGAAAGACCA 1500	
Db	1702 GTCATCATGACCAAAATGAGAACATTCAGTGGTGCAGAAATAGGCTGCGAAGAAAGACCA 1761	
Qy	1501 GAAACACATCAACCAATCAGAGACAGAAATGCTTTAAACATTCATCTAAACATATAGCTTGCCT 1560	
Db	1762 GAAACACATCAACCAATCAGAGACAGAAATGCTTTAAACATTCATCTAAACATATAGCTTGCCT 1821	
Qy	1561 GGTGACTTTGGTTGGTTGTCACAAAATGAATGGCCCATGATTTGCTGGTTGGTAAAG 1620	
Db	1822 GGTGACTTTGGTTGGTTGTCACAAAATGAATGGCCCATGATTTGCTGGTTGGTAAAG 1881	
Qy	1621 AATGGTTACCAATCTACCATGGCAAGCTACTGCTTAATGGGGCAAAAGTTCTCTGATTGG 1680	
Db	1882 AATGGTTACCAATCTACCATGGCAAGCTACTGCTTAATGGGGCAAAAGTTCTCTGATTGG 1941	
Qy	1681 TCAGAAAACTGGCGGAGCCAAAGGTGCCAACTCTTATATAATTTACTAGGTTTCGGCAGCG 1740	
Db	1942 TCAGAAAACTGGCGGAGCCAAAGGTGCCCACTCTTATATAATTTACTAGGTTTCGGCAGCG 2001	
Qy	1741 TCACCAATTCAGACACCGAAAAAGTACGCTCTCAGCCAGAACTATGCACTTAATCTCACTT 1800	
Db	2002 TCACCAATTCAGACACCGAAAAAGTACGCTCTCAGCCAGAACTATGCAATAAATCTCACTT 2061	
Qy	1801 GCATCGATCTCGAGACCTTGGCTTTAGAGCTTTGGAGCACAACCAATACTCTCTGTTGG 1860	
Db	2062 GCATCGATCTCGAGACCTTGGCTTTAGAGCTTTGGAGCACAACCAATACTCTCTGTTGG 2121	
Qy	1861 GGCACTGCAAGAAACCCAGAACTCTGGGAGAGCTGGTTTCCAAAGCCCTGCCAAGATGCTCAA 1920	
Db	2122 GGCACTGCAAGAAACCCAGAACTCTGGGAGAGCTGGTTTCCAAAGCCCTGCCAAGATGCTCAA 2181	
Qy	1921 CTGAGCCCAACTTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTGGTGGGAACCG 1980	
Db	2182 CTGAGCCCAACTTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTGGTGGGAACCG 2241	
Qy	1981 TTGAAAGAAAGACTTTGAGCGAGCGCTTGAATTCGACTAA 2019	
Db	2242 TTGAAAGAAAGACTTTGAGCGAGCGCTTGAATTCGACTAA 2280	

```
DEFINITION Mouse parvovirus 1b DNA.
ACCESSION U34253
VERSION U34253.1 GI:1464793
KEYWORDS
SOURCE Mouse parvovirus 1b
ORGANISM Mouse parvovirus 1b
REFERENCE 1 (bases 1 to 4764)
AUTHORS Besselsen, D.G., Pintel, D.J., Purdy, G.A., Besch-Williford, C.L.,
Franklin, C.L., Hook, R.R. Jr. and Riley, L.K.
TITLE Molecular characterization of newly recognized rodent parvoviruses
JOURNAL J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
MEDLINE 96201434
PUBMED 8609486
REFERENCE 2 (bases 1 to 4764)
AUTHORS Besselsen, D.G.
TITLE Direct Submission
JOURNAL Submitted (17-AUG-1995) David G. Besselsen, Department of
Veterinary Pathology, University of Missouri-Columbia, W213
Veterinary Medicine Building, Columbia, MO 65211, USA
FEATURES
source
1..4764
/organism="Mouse parvovirus 1b"
/mol_type="genomic DNA"
/species_host="Mus musculus"
/db_xref="taxon:42841"
ORIGIN
Query Match 94.2%; Score 1902.2; DB 14; Length 4764;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 1946; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
Qy 1 ATGGCTGGAATGCTTACTCTGATGAAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 60
Db 122 ATGGCTGGAATGCTTACTCTGATGAAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 181
Qy 61 AGTAACCGAGGAGTGTCTCATTTGTTTTTAAAGTAAAGTGTCACTGAATGGAAGA 120
Db 182 AGTAACCGAGGAGTGTCTCATTTGTTTTTAAAGTAAAGTGTCACTGAATGGAAGA 241
Qy 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGAGGAGGAGGAGCTGGAATCTTTACA 180
Db 242 GATATCGGATGGAATAGTTTACAAAAAGAGCTGAGGAGGAGGAGCTGGAATCTTTACA 301
Qy 181 CGAGGAGCGGAACTACTTGGGACCAAGCGAGGACATGGAATGGGAAACCACTGGAT 240
Db 302 CGAGGAGCGGAACTACTTGGGACCAAGCGAGGACATGGAATGGGAACTCTACAGT 361
Qy 241 GAAATGACCAAAAAAGCAAGTATTCATTTTGGTATTTGTTTAAAGTGTATTTGAA 300
Db 362 GAAATGACCAAAAAAGCAAGTATTCATTTTGGTATTTGTTTAAAGTGTATTTGAA 421
Qy 301 GTGCTTAACCAAGAAATATATTTCTGGTGATGTTAAATGGTTGTCACATGATGG 360
Db 422 GTGCTTAACCAAGAAATATAGCTCTCTGCTGATGTTTACTTGGTTTGGCAGCAT 481
Qy 361 GGAAGAAGCAAGGCTGGCACTGCTACTTAATTTGGAGGAAGGAGCTTTAGTCAAGCT 420
Db 482 GGAAGAAGCAAGGCTGGCACTGCTACTTAATTTGGAGGAAGGAGCTTTAGTCAAGCT 541
Qy 421 CAAGGGAATGGTGGAGAGGCACTAAATGTTTACTTGGAGGAGATGGTTGGTAAAC 480
Db 542 CAAGGGAATGGTGGAGAGGCACTAAATGTTTACTTGGAGGAGATGGTTGGTAAAC 601
Qy 481 TGTAAATGCAACTTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT 540
Db 602 TGTAAATGCAACTTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT 661
Qy 541 GAGTGGGTACTCTACTTACTTAAAGCATAAGCAAAACCAAAAAAGACTTACCAAGTGT 600
Db 662 GAGTGGGTACTCTACTTACTTAAAGCATAAGCAAAACCAAAAAAGACTTACCAAGTGT 721
Qy 601 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAAAATAGCACTAGT 660
```

```
Db 1802 ACAGAAACTGGCGGAGCCGAGGTGCGACTCCTATATAATTCACCTAGGTTCGGCAGCG 1861
Qy 1741 TCACCATTCACGACACCGAAAGTACGCTCTCAGCAGAACATATGCACTAACCTCCACTT 1800
Db 1862 TCACCATTCACGACACCGAAAGTACGCTCTCAGCAGAACATATGCACTAACCTCCACTT 1921
Qy 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACCACCAATATCTCTGTTCGG 1860
Db 1922 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACCACCAATATCTCTGTTCGG 1981
Qy 1861 GGCACTCGAGAAACCCAGAACACTGGGAAAGCTGGTTCCAAAGCCTCCCAAGATGGTCAA 1920
Db 1982 GGCACTCGAGAAACCCAGAACACTGGGAAAGCTGGTTCCAAAGCCTCCCAAGATGGTCAA 2041
Qy 1921 CTGAGCCCAACTTGGTCAGAGATCAGAGAGATTTGAGACGCTGCTTCGGTGGGAAACCG 1980
Db 2042 CTGAGCCCAACTTGGTCAGAGATCAGAGAGATTTGAGACGCTGCTTCGGTGGGAAACCG 2101
Qy 1981 TTGAAGAAAGACTTCAGGAGCGCGCTGAACCTTGACTAA 2019
Db 2102 TTGAAGAGAGACTTCAGGAGCGCGCTGAACCTTGACTAA 2140

RESULT 10
PAMVMI
LOCUS Mouse parvovirus minute virus immunosuppressive variant genome (= 5087 bp DNA linear VRL 01-JUL-1999
DEFINITION MWi).
ACCESSION X02481.1 GI:60918
VERSION coat protein; genome; origin of replication; overlapping genes;
KEYWORDS terminal repeat; unidentified reading frame.
SOURCE Mouse minute virus
ORGANISM Mouse minute virus
VIRUSES: ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE 1 (bases 1 to 5087)
AUTHORS Sahli,R., McMaster,G.K. and Hirt,B.
TITLE DNA sequence comparison between two tissue-specific variants of the
autonomous parvovirus, minute virus of mice
JOURNAL Nucleic Acids Res. 13 (10), 3617-3633 (1985)
MEDLINE 85242059
PUBMED 3855242
COMMENT For the fibroblast-specific strain (MWmp) sequence see <PAMW2>.
The genomes of MWmp and MWmi (immunosuppressive variant) have more
than 96% of their sequence in common.
Data kindly reviewed (18-JUL-1986) by G. McMaster.

FEATURES
source 1..5087
/organism="Mouse minute virus"
/mol_type="genomic DNA"
/db_xref="taxon:10794"
misc_feature 1..116
/notes="terminal hairpin region"
promoter 176..183
/notes="TATA box"
gene 2405..4558
/genes="VP1"
CDS <2405..4558
/genes="VP1"
/function="coat protein"
/codon_start=1
/product="vp1 protein"
/protein_id="CAB46507.1"
/db_xref="GI:5419928"
/db_xref="GOA:P07302"
/db_xref="Swiss-Prot:P07302"
/translation="VPGYKYLPGNSLDQGEPTNPSDAAKSHDBAYDQYIKSGKNP
YLFSADQRFIDQTKADKWGGKVGHYFFRTKRAFAPKLATDSEFGTSGVRAGRT
RPPAYIFINQARAKLTSAQSSQTSMDTSGTQPDGNGVHSAARVERAADPGGS
GGGGGGGGVSTGSDYNQTHYRFLGDGWEITATALRLVHLNMPKSENYCRIRVHN
TTDTSVKGKNAKDDAHEQIWPMSLDANAWGVLQPSDWQYICNTMSQLNLSLQDE
IFNVLKTVEQDSGGQAIKIYNNDLTACMMVAVDNNILPYTPAANSMETLGFYFWK
PTIASPYRYFYFCVDRDLSTVYENQEGTIEHNVTGTPKGMNSQFTTIENTQQITLLRTG
```

```
DEPATCTYYFDTPNPVKLTHTWQTNROLGPPLLSTPEADTDAGTLTAQSRHGATQM
EVNWSEAIRTPAQVGFQPHNDPEASRAGPFAAPKVPADVTQPPQVPPQVREANGSVRSY
GKQVGNWAAHGAPAPERYTWDETFNFGSGDRDTRDGTIOAPLVPVPPGLNGLTNANPIG
TKNDIHFSNFSYGLPTAFSHSPVYPOQIWDKELDLHKPLHITAFVCKNNAP
GOMLVRGLNLTQDYPNGATLSRIVTYGTFFWKGLTMRKLRANTTNWNPVYQVSVE
DNGNSYMSVTKMLPTATGNMQSVPLITRPVARTY"
2792..4558
/genes="VP2"
<2792..4558
/genes="VP2"
/function="coat protein"
/codon_start=1
/product="VP2 protein"
/protein_id="CAB46508.1"
/db_xref="GI:5419929"
/db_xref="GOA:P07302"
/db_xref="Swiss-Prot:P07302"
/translation="TMSDGTSDPDGNGVHSAARVERAADPGSGGGGGGGGGVGS
TGSYDQTHYRFLGDGWEITATALRLVHLNMPKSENYCRIRVHTTDTSVKGNMAKD
DAHEQIWPMSLDANAWGVLQPSDWQYICNTMSQLNLSLQDEIFNVWVKTVTEQD
SGQAIKIYNNDLTACMMVAVDNNILPYTPAANSMETLGFYFWKPTIASPYRYFCV
DRDLSTVYENQEGTIEHNVTGTPKGMNSQFTTIENTQQITLLRTGDEFFATGYTFTDN
PVKLTHTWQTNROLGPPLLSTPEADTDAGTLTAQSRHGATQMEVNWSEAIRTRP
AOVGFCOPHNDPEASRAGPFAAPKVPADVTQGVDRANGSVRSYKQHGKNAHGP
APERYTWDETFNFGSGDRDTRDGTIOAPLVPVPPGLNGLTNANPITKNDLHFSNVNS
YGLPTAFSHSPVYPOQIWDKELDLHKPLHITAFVCKNNAPGOMLVRGLNLTLD
QYDYPNGATLSRIVTYGTFFWKGLTMRKLRANTTNWNPVYQVSVEDNGNSYMSVTKWL
PTATGNMQSVPLITRPVARTY"
4603..4608
/misc_feature /note="polyadenylation signal"
4821..4826
/misc_feature /note="polyadenylation signal"
4878..5087
/misc_feature /note="terminal hairpin region"

ORIGIN
Query Match 94.2%; Score 1902.2; DB 14; Length 5087;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 1946; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
Qy 1 ATGGCTGGAAATGCTTACTCTGATGAAGTGTTCGGAGCAACCACTGGTTAAAGGAAAA 60
Db 262 ATGGCTGGAAATGCTTACTCTGATGAAGTGTTCGGAAACAACCACTGGTTAAAGAAAA 321
Qy 61 AGTAACCAAGAGTGTCTCATTTGTTTAAATAATGAAATGTTCAACTGAATGGAAAA 120
Db 322 AGTAACCAAGAGTGTCTCATTTGTTTAAATAATGAGGATGTTCAACTGAATGGAAAA 381
Qy 121 GATATCGGATGGAATAGTTACAAAAAGAGCTCGAGGAGACGAGCTGAATCTTTACAA 180
Db 382 GATATCGGATGGAATAGTTACAAAAAGAGCTCGAGGAGACGAGCTGAATCTTTACAA 441
Qy 181 CGAGGAGCGAATACTACTTGGGACCAAGCAGGACATGGAATGGGAAACACAGTGGAT 240
Db 442 CGAGGAGCGAATACTACTTGGGACCAAGCAGGACATGGAATGGGAAATCTACAGTGGAT 501
Qy 241 GAAATGACCAAAAAAGCAAGTATTTCATTTTGTGTTTAAATAATGTTTATTTGAA 300
Db 502 GAAATGACCAAAAAAGCAAGTATTTCATTTATGACTCTTATGTTAAATAATGTTTGTGAA 561
Qy 301 GTGCTTAAACAAAAAGTATATATTTCTCTGTTGATGTTAAATTTGTTGTGCAACATGATGG 360
Db 562 GTGCTTAGCACAAAAAATATAGCTCTCTGCTGATGTTACTTGTGTTTGTGACAGCATGATGG 621
Qy 361 GGAAGAGCAACGCTGGCAGCTGCCACTGACTAATTTGGAGGAAGGAGCTTTAGTCAAGCT 420
Db 622 GGAAGAGCAACGCTGGCAGCTGCCACTGACTAATTTGGAGGAAGGAGCTTTAGTCAAGCT 681
Qy 421 CAAGGAAATGTTGGAGAAAGGCAATTAATGTTTACTTGGAGCAGATGTTGTTGTAACAGCC 480
Db 682 CAAGGAAATGTTGGAGAAAGGCAATTAATGTTTACTTGGAGCAGATGTTGTTGTAACAGCC 741
Qy 481 TGTAATGTGCAACTAACACACGCTGAAAGAAATTAATAAGAGAAATAGCAGAAGCAAT 540
```

```
Db      742  TGTAAATGTCAGCTAAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAGACAGT 801
Qy      541  GAGTGGGTACTCTACTTACTTATTAAGCATTAAGCAAAACCAAAAGAGACTATACCAAGTGT 600
Db      802  GAGTGGGTACTCTACTTACTTATTAAGCATTAAGCAAAACCAAAAGAGACTATACCAAGTGT 861
Qy      601  GTTCTTTTGGAAACATGATGCTTACTTATTTTAACTTAAAGAAAGAAATAAGCACTAGT 660
Db      862  GTTCTTTTGGAAATATGATGCTTACTTATTTTAACTTAAAGAAAGAAATAAGCACTAGT 921
Qy      661  CCACAAGAGACGAGGCTATTTTCTTAGCAGTCACTCTGGCTGGAAACCTAACTTTTAA 720
Db      922  CCGCCAGGACGAGGCTATTTTCTTAGCAGTCACTCTGGCTGGAAACCTAACTTTTAA 981
Qy      721  AAAGAACGCGAGCCCATCTAGTGGCAAACTATACACTGATGATGATGATGATGATGATGAT 780
Db      982  AAAGAGGCGCAACGCGCATCTAGTGGCAAACTATACACTGATGATGATGATGATGATGATGAT 1041
Qy      781  GTTGAACACACAGTAACCACTGCGCAGGAACTAAGCGCGGCGAGAACTTCAAACTTAAAAA 840
Db      1042  GTTGAACACACAGTAACCACTGCGCAGGAACTAAGCGCGGCGAGAACTTCAAACTTAAAAA 1101
Qy      841  GAACTTTCTATTAATACTAATTAAGAGCTGTGATTAAGAGTAACCTACACAGAG 900
Db      1102  GAGGTTTCTATTAATACTAATTAAGAGCTGTGATTAAGAGTAACCTACACAGAG 1161
Qy      901  GACTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Db      1162  GACTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1221
Qy      961  AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAATCTAGCAGAACCAAAACAGCA 1020
Db      1222  AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAATCTAGCAGAACCAAAACAGCA 1281
Qy      1021  TTTGACTTAATTTTAAAGAAAGCTGAAACCAAGCAAACTAAGCAAACTTCACTGCTGAC 1080
Db      1282  TTTGACTTAATTTTAAAGAAAGCTGAAACCAAGCAAACTAAGCAAACTTCACTGCTGAC 1341
Qy      1081  ACAAGAACCTGCAAGATTTTGTCTTTTATGCTGCTGGAATATGTTAAAGTTGCGCATGCT 1140
Db      1342  ACAAGAACCTGCAAGATTTTGTCTTTTATGCTGCTGGAATATGTTAAAGTTGCGCATGCT 1401
Qy      1141  ATTTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATGCTGTTTATTTTATTCAGACCA 1200
Db      1402  ATTTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATGCTGTTTATTTTATTCAGACCA 1461
Qy      1201  GCCAGCACAGCAATCTATTTATGCAAGCCATAGCACAAGCAGTGGCAATGTTGGT 1260
Db      1462  GCCAGTACAGCAATCTATTTATGCAAGCCATAGCACAAGCAGTGGCAATGTTGGT 1521
Qy      1261  TGCTATAATGAGCAATGTAACCTTTTCCATTTAATGACTGTACCAACAGAACTTGATT 1320
Db      1522  TGCTATAATGAGCAATGTAACCTTTTCCATTTAATGACTGTACCAACAGAACTTGATT 1581
Qy      1321  TGGGTAGAAGAGCTGTTAACTTTTGGACAGCAAGTAACAGTTTAAAGCAATTTGCTCT 1380
Db      1582  TGGGTAGAAGAGCTGTTAACTTTTGGACAGCAAGTAACAGTTTAAAGCAATTTGCTCT 1641
Qy      1381  GGTCAAACTATTCGATTTGATCAAAAGAAAGGAGGAGCAAAACAGTTGAACCAACCA 1440
Db      1642  GGTCAAACTATTCGATTTGATCAAAAGAAAGGAGGAGCAAAACAGTTGAACCAACCA 1701
Qy      1441  GTCAATCATGACCAAAATGAGAACATTAAGTGTGATGATGATGATGATGATGATGATGATGAT 1500
Db      1702  GTCAATCATGACCAAAATGAGAACATTAAGTGTGATGATGATGATGATGATGATGATGATGAT 1761
Qy      1501  GAACACACTCAACCAATCAGACAGAGATGCTTAACTTCACTTAAACACATACCTTGCCT 1560
Db      1762  GAACACACTCAACCAATCAGACAGAGATGCTTAACTTCACTTAAACACATACCTTGCCT 1821
Qy      1561  GGTGACTTTGGTTGGTTGACAAATGAATGCGCCCATGATTTGCTTGGTTGGTAAAG 1620
```

```
Db      1822  GGTGACTTTGGTTGGTTGACAAAGTAATGAGCCCATGATTTGCTTGGTTGGTAAAG 1881
Qy      1621  AATGGTTTACCAATCTACCACTGGCAAGCTACTGCTTAAATGGGGCAAGTTCTCTGATTGG 1880
Db      1882  AATGGTTTACCAATCTACCACTGGCAAGCTACTGCTTAAATGGGGCAAGTTCTCTGATTGG 1941
Qy      1681  TCAGAAAACTGGCGGAGCCAAAGGTGCGAACTCTCTATAAATTTACTAGGTTCGGACGCG 1740
Db      1942  TCAGAAAACTGGCGGAGCCAAAGGTGCGAACTCTCTATAAATTTACTAGGTTCGGACGCG 2001
Qy      1741  TCACCAATTCAGACACCGAAAAAGTACGCTCTCAGCCAGAACTATGCACTTAACTCCACTT 1800
Db      2002  TCACCAATTCAGACACCGAAAAAGTACGCTCTCAGCCAGAACTATGCACTTAACTCCACTT 2061
Qy      1801  GCATCGATCTCGAGACCTGCGCTTTAGAGCCTTGGAGCCACCAATTAATCTCTCTGTTGG 1860
Db      2062  GCATCGATCTCGAGACCTGCGCTTTAGAGCCTTGGAGCCACCAATTAATCTCTCTGTTGG 2121
Qy      1861  GGCACCTGCAGAAACCCAGAACACTGCGGGAAGCTGGTTCCAAAGCCTGCCAAGATGCTCAA 1920
Db      2122  GGCACCTGCAGAAACCCAGAACACTGCGGGAAGCTGGTTCCAAAGCCTGCCAAGATGCTCAA 2181
Qy      1921  CTGAGCCCAACTTGGTTCAGAGATCGAGAGGATTTGAGAGCGTGTTCGGTTCGGGAACCG 1980
Db      2182  CTGAGCCCAACTTGGTTCAGAGATCGAGAGGATTTGAGAGCGTGTTCGGTTCGGGAACCG 2241
Qy      1981  TTCAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019
Db      2242  TTCAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2280
```

RESULT 11

```
MPU12469
LOCUS      MPU12469
DEFINITION Mouse parvovirus 1 complete genome, nonstructural protein 1 (NS1)
            gene, complete cds and capsid protein (VP1) gene, complete cds.
ACCESSION U12469
VERSION   U12469.1 GI:525325
KEYWORDS  Mouse parvovirus 1
SOURCE    Mouse parvovirus 1
ORGANISM  Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE 1 (bases 1 to 5144)
AUTHORS   Ball-Goodrich,L.J. and Johnson,E.
TITLE     Molecular characterization of a newly recognized mouse parvovirus
JOURNAL   J. Virol. 68 (10), 6476-6486 (1994)
MEDLINE   9435951
PUBMED    8083985
REFERENCE 2 (bases 1 to 5144)
AUTHORS   Ball-Goodrich,L.J.
TITLE     Direct Submission
JOURNAL   Submitted (20-JUL-1994) Lisa J. Ball-Goodrich, Section of
            Comparative Medicine, Yale University School of Medicine, New
            Haven, CT 06520-8016, USA
```

FEATURES

```
source      Location/Qualifiers
            1..5144
            /organism="Mouse parvovirus 1"
            /mol_type="genomic DNA"
            /db_xref="taxon:35340"
            /clone="PRVXB36; plutB/R13-6; pRTBg/H3-1"
            265..2283
            /gene="NS1"
            265..2283
            /gene="NS1"
            /codon_start=1
            /product="nonstructural protein 1"
            /protein_id="AAA61405.1"
            /db_xref="GI:525326"
            /translation="MAGNAYSDEVILGTTNWLKEKSNQEVFSFVKTEVDVQNGKDITGW
            NNYKELODELKSLQGAETTWDSDEMEWETVDEMTKQVFIYDSLVRKCLPEVL
            STNTAPADVTWFVQHEKDKDQWCHVLIIGKDFSOAQGWKRWRLQNVYWRMLVTA
            CNVQLTPABRIKLREIADSESWTLLTYKHQTKDKDYKCVLFGNMLIAYFLTKKKS
            TSPPRDGGYFUSSDSGWKTNFKGERHLVSKLYITDMDRPETVETTTVTTAQTKRGRI
```

```
OTKKEYSIKTTIKELVHKRVTSBEDWMMQPDYSIEMMAQPGGNNLKNLTLEICTLTL
ARTKAPDLLEKAETSKLNFSLPDRITCKYIFAFHGWNYIKVCHACCVLNRQGGKR
NTVLFPGPASTGSIIIAQAQVAGNVGVCYKAAIVNFDFNDCTNKLNIWEEAGNFGQ
QVNFQKAI CSQGTIRIDQKGGKQIEPTVIMTINENITVVRIGCEERPEHTQPIRD
RMLNHLTHLPDGLVDKNWPMICAWLKGXGYOSTWASYCAKRGKVPDHTQPNWAE
PKVPTPI NSIGSARSPTTPKSP LSONYALTP LADSLDLELALPEHSTENTPVAGTAE
TQNTGAGSKACQDQLSP TWSIEBDRACFGAEPLKKDFSEPLND"
/gene="vpi"
join(2290..2320,2403..4561)
CDS
/join(2290..2320,2403..4561)
/gene="vpi"
/codon_start=1
/product="capsid protein"
/protein_id="AA61406.1"
/db_xref="GI:525327"
/tranlations="MAPPAKRAKRGWVPYGYKYLGPNSLDQGEPTNPDAAKEHDE
AYDKYIKSGNPLYFSAADORFI DTKADGWGKVYFFRTKFAFAPRLASSSEP
GTSVGSITAGKRTKPAHIFINQARAKKRASLAQOORTOTMSDGAQPOSGSAVQSA
RVERAADPGSGGGSGGGGVGVSQYNDQTHYFLSDGWVEITATYSTRVHLNMP
KSENYCRVRVHNTNRTAGNMAKDDAHEQIWTPLSDNSNAGWVFPQSDWQFINN
MSHVNLSLDOELFNVIKTVTEQNTGAERAKVIYNNDLTASMMVALDSNNILPYTAT
DNQETIGFYFWKTPSPRYRYFNCDSLSVTYDQTSIVDTMANASGLSOPFFIE
NTQRIOLLRTGDSFAGCTYFETPEPIKLSHTWQSNQLGQPOPI TDLPTADNENATLV
TRGDRSGITQISGSDNVTEATRVBAQVGFQCPHDNFEETSRAGPKVVPVNAVYTOGN
EHDANGSLRYTDKQGDWGSNNKSEKFTWDALSYDSGRWADRCFINATPFTSPPAL
NHLITNSDPIGNKTAIHYQNVFNSYGLPTAFPHAPPIYPOQIWDKELDLHKPRLHA
QAPFVCKNNAAPGLLRLAPNLTDQYD PMSSTLSRIYTYGTFPFWKGLTKLAKLRPNA
TWNPVYQVSAQYQNEENEYSIHKWLPTATGNMQSIPLLSRPVARNTY"
```

ORIGIN

```
Query Match 94.1%; Score 1899; DB 14; Length 5144;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1944; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 1 ATGCTGGAATGCTTACTCTGATGAAGTTTTCGGAGCAACCACTGGTTTAAAGGAAAA 60
DB 265 ATGCTGGAATGCTTACTCTGATGAAGTTTTCGGAGCAACCACTGGTTTAAAGGAAAA 324

QY 61 AGTAACGAGGAAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAAATGGAATA 120
DB 325 AGTAACGAGGAAGTGTCTCATTTGTTTTTAAAACTGAGGATGTTCAACTAAATGGAATA 384

QY 121 GATATCGGATGGAATAGTTTACAAAAGAGCTCGAGGAGCAAGCTGGAATCTTTACAA 180
DB 385 GATATCGGATGGAATTAATACAAAAGAGCTCGAGGAGCAAGCTGGAATCTTTACAA 444

QY 181 CGAGGAGCGAAACTACTTCGGACCAAGCGAGGACATGGAATGGGAAACACAGTGGAT 240
DB 445 CGAGGAGCGAAACTACTTCGGACCAAGCGAGGACATGGAATGGGAAATCTACAGTGGAT 504

QY 241 GAAATGACCAAAAAGCAAGTATTTCATTTTGTATTCTTTGTTTAAAAAATGTTTATTGAA 300
DB 505 GAAATGACCAAAAAGCAAGTATTTCATTTATGACTCTTTAGTTAAAAAATGTTTATTGAA 564

QY 301 GTGCTTAA CACAAAGAAATATTTTCTGTGTGATGTTAAATTTGTTTGTGCAACATGAATGG 360
DB 565 GTGCTTAGCA CAAAGAAATATAGCTCCTGCTGATGTTACTTGTGTTTGTGAGCATGAATGG 624

QY 361 GGAAGAGACCAAGCTGGCACTCCCATGACTAATTTGGAGGAAGGACTTTTACTCAAGCT 420
DB 625 GGAAGAGACCAAGCTGGCACTGCCATGTACTAATTTGGAGGCAAGGACTTTTACTCAAGCT 684

QY 421 CAAGGGAATATGTTGGAGAGCAAGCAATTAATTTTACTTGGAGCAGATGTTGGTAAACAGCC 480
DB 685 CAAGGGAATATGTTGGAGAGCAAGCTAATTTTACTTGGAGCAGATGTTGGTAAACAGCC 744

QY 481 TGTAAATGTCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGAACAT 540
DB 745 TGTAAATGTGAGCTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGACAGT 804

QY 541 GAGTGGGTACTTACTTACTTATAGCATAAGCAACCAAAAAGACTATATACAGATGT 600
DB 805 GAGTGGGTACTTACTTACTTATAAACATTAAGCAAAACCAAAAAGGACTATATACATGT 864
```

```
QY 601 GTTCTTTTGGAAACAATGATTTGCTTACTATTTTAACTAAAAAGAAAAATAAGCACTAGT 660
DB 865 GTTCTTTTGGAAATATGATTTGCTTACTATTTTAAACCAAAAAAATAAGCACCAGT 924

QY 661 CCACCAAGAGACGGAGGCTATTTTCTTACGAGTGACTCTGGCTGGAAAACTTAATCTTTTA 720
DB 925 CGCCCAAGAGACGGAGGCTATTTTCTTAAAGCAGTGACTCTGGCTGGAAAACTTAATCTTTTA 984

QY 721 AAAGAAGGCGAGCGCCCATCTAGTGAGCAAACTATACACTGATGACATCGCGCCGAGAAACG 780
DB 985 AAAGAAGGCGAGCGCCCATCTAGTGAGCAAACTATACACTGATGACATCGCAACGAAACG 1044

QY 781 GTTGAACCAACAGTAAACCACTGCGCAGGAAACTTAAAGCGGCGAGAAATTTCAAACTAAAAA 840
DB 1045 GTTGAACCAACAGTAAACCACTGCGCAGGAAACTTAAAGCGGCGAGAAATTTCAAACTAAAAA 1104

QY 841 GAAGTTTCTATTAAAACTACCTTAAAGAGCTGGTGCATAAAAAGAGTAACTCACCAGAG 900
DB 1105 GAGTTTCTATTAAAAACCACTTAAAGAGCTGGTGCATAAAAAGAGTAACTCACCAGAA 1164

QY 901 GACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 1165 GACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1224

QY 961 AACTCTGCTGAAAAAATACGCTAGAGATTTGTACACTAATCTAGCCAGAACCAAAACAGCA 1020
DB 1225 AACTCTGCTGAAAAAATACGCTAGAGATCTGTACACTAATCTAGCCAGAACCAAAACAGCA 1284

QY 1021 TTTGACTTAAATTTTAGAAAAAGCTGAACCAAGCAAACTAACCACTTTTCACTGCTGCTGAC 1080
DB 1285 TTTGACTTAAATTTTAGAAAAAGCTGAACCAAGCAAACTAACCACTTTTCACTGCTGCTGAC 1344

QY 1081 ACAAGAACCTGCAAGATTTTGTCTTTTTCATGGCTGGAATCTATCTTAAAGTTTTCGCCATGCT 1140
DB 1345 ACAGAACCTGCAAGATCTTTGTCTTTTTCATGGCTGGAATCTATCAATTAAGTTTTCGCCATGCT 1404

QY 1141 ATTGCTGTGTTTTTAAACAGACAAAGAGGCAAAAGAAATGCTGTTTTTATTTTCAATGAGCA 1200
DB 1405 ATTGCTGTGTTTTTAAACAGACAAAGAGGCAAAAGAAATGCTGTTTTTATTTTCAATGAGCA 1464

QY 1201 GCCAGACAGGCAAAATCTATTATTGCAAGCCATAGCACAGGCTTGGCAATGTTGGT 1260
DB 1465 GCCAGTACAGGCAAAATCTATTATTGCAAGCCATAGCACAGGCTTGGTAAATGTTGGT 1524

QY 1261 TGCTATAATGCAAGCAATGTAAACTTTTCCATTTTAAATGACTGTACCAACAGAACTTGATT 1320
DB 1525 TGCTATAATGCAAGCAATGTAAACTTTTCCATTTCAATGACTGTACCAACAGAACTTAAAT 1584

QY 1321 TGGGTAGAAAGAGCTGGTAACTTTTGGACAGCAAGTAAACCAAGTTTAAAGCCATTGCTCT 1380
DB 1585 TGGGTAGAAAGAGCTGGTAACTTTTGGACAAACAGTAAACCAAGTTTAAAGCCATTGCTCT 1644

QY 1381 GGTCAAACTATTTCGCAATTGATCAAAAAGGAAAGGAGCAACACAGATTGAAACCAACCA 1440
DB 1645 GGTCAAACTATTTCGCAATTGATCAAAAAGGAAAGGAGCAACAGATTGAAACCAACCA 1704

QY 1441 GTCATCATGACCAAAATGAGAACATTTACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA 1500
DB 1705 GTCATCATGACCAAAATGAGAACATTTACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA 1764

QY 1501 GAAACACATCAACCAATTCAGAGACAGAAATGAAATGGCCCATGATTTGCTGTTGGTAAAG 1560
DB 1765 GAAACACATCAACCAATTCAGAGACAGAAATGAAATGGCCCATGATTTGCTGTTGGTAAAG 1824

QY 1561 GGTGACTTTGGTTGGTTGACAAAAATGAAATGGCCCATGATTTGCTGTTGGTAAAG 1620
DB 1825 GGTGACTTTGGTTGGTTGACAAAAATGAAATGGCCCATGATTTGCTGTTGGTAAAG 1884

QY 1621 AATGGTTTCAATCTACCATGCGCAAGCTACTGTCTAAATGGGGCAAGTTCTCTGATTGG 1680
DB 1885 AATGGTTTCAATCTACCATGCGCAAGCTACTGTCTAAATGGGGCAAGTTCTCTGATTGG 1944
```



```
QY 1681 TCAGAAACTGGCGGAGCCAAAGTGCCTTAAATTTCTAGTTCGGCAGC 1740
Db 1945 ACGGAAACTGGCGGAGCCGAGTGCCTTAAATTTCTAGTTCGGCAGC 2004
QY 1741 TCACCATTCAGACACCGAAAGTACGCTCTAGCCAGAACATGCACTT 1800
Db 2005 TCACCATTCAGACACCGAAAGTACGCTCTAGCCAGAACATGCACTT 2064
QY 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGGAGCACACCAATATCTCTGTTGCG 1860
Db 2065 GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGGAGCACACCAATATCTCTGTTGCG 2124
QY 1861 GGCACTCGAGAAACCCAGAACACCTGGGAAAGCTGTTCCAAAGCTCCCAAGATGTCFAA 1920
Db 2125 GGCACTCGAGAAACCCAGAACACCTGGGAAAGCTGTTCCAAAGCTCCCAAGATGTCFAA 2184
QY 1921 CTGAGCCCACTTGGTTCAGAGATCGAGGAGATTTGAGAGGCTGCTTCGGTGGGAACCG 1980
Db 2185 CTGAGCCCACTTGGTTCAGAGATCGAGGAGATTTGAGAGGCTGCTTCGGTGGGAACCG 2244
QY 1981 TTGAAGAAAGACTTTCAGCGAGCCCTGAACCTTGACTAA 2019
Db 2245 TTGAAGAAAGACTTTCAGCGAGCCCTGAACCTTGACTAA 2283

RESULT 12
MOU34254
LOCUS
DEFINITION Mouse parvovirus 1c DNA.
ACCESSION U34254
VERSION U34254.1 GI:1464794
KEYWORDS
SOURCE
ORGANISM
Mouse parvovirus 1c
Mouse parvovirus 1c
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE
1 (bases 1 to 4764)
Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L.,
Franklin,C.L., Hook,R.R. Jr. and Riley,L.K.
Molecular characterization of newly recognized rodent parvoviruses
J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
MEDLINE
96201434
FUBMED
8609486
REFERENCE
2 (bases 1 to 4764)
Besselsen,D.G.
Direct Submission
Submitted (17-AUG-1995) David G. Besselsen, Department of
Veterinary Pathology, University of Missouri-Columbia, W213
Veterinary Medicine Building, Columbia, MO 65211, USA
FEATURES
source
Location/Qualifiers
1..4764
/organism="Mouse parvovirus 1c"
/mol_type="genomic DNA"
/specific_host="Mus musculus"
/db_xref="taxon:42842"

ORIGIN
Query Match 92.8%; Score 1873; DB 14; Length 4764;
Best Local Similarity 95.4%; Pred. No. 0;
Matches 1927; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGAGCAACCAACTGGTTAAAGGAAAA 60
Db 122 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGCAACCAACCACTGGTTAAAGGAAAA 181
QY 61 AGTAACGAGAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGAAAA 120
Db 182 AGTAACGAGAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGAAAA 241
QY 121 GATATCGGATGGAATAGTTTCAAAAAAGAGCTGCGAGGAGCAGAGCTGGAATCTTTACAA 180
Db 242 GATATCGGATGGAATANTTACAGNAGGAGCTGCAAGAGAGCAGGCTAAATCTTTACAA 301
QY 181 CGAGGAGCGGAACTACTTGGGACCAAGCGAGGACATGGAATGGGAAACCAACAGTGGAT 240
```

```
Db 302 CGAGGAGCGGAAACTACTCTGGGACCAAGCGAGGACATGGAATGGGAATCTACAGTGGAT 361
QY 241 GAATGACCAAAAGCAAGTATTCTTTTCTTTTGGTTAAATAATGTTTATTGAA 300
Db 362 GAAGTGACCAAAAGCAAGTATTCTTTTATGACTCTTTTAGTTAAATAATGTTTATTGAA 421
QY 301 GTCTTTAAACAACAAAGAAATATATTTCTGGTGATGTTAAATTTGGTTTCTGCAACATGAATGG 360
Db 422 GTCTTTAAACAACAAAGAAATATATTTCTGGTGATGTTAAATTTGGTTTCTGCAACATGAATGG 481
QY 361 GGAAGAAGCAAGGCTGGCACTGCAATGTAATAATTTGGAGGAAAGCACTTTAGTCAAGCT 420
Db 482 GGAAGAAGCAAGGCTGGCACTGCAATGTAATAATTTGGAGGAAAGCACTTTAGTCAAGCT 541
QY 421 CAAAGGAAATGTTGGAGAGGCACTAAATGTTTACTGGAGCAGATGTTGGTAAACAGCC 480
Db 542 CAAAGGAAATGTTGGAGAGGCACTAAATGTTTACTGGAGCAGATGTTGGTAAACAGCC 601
QY 481 TGTAAATGTGCAACTAAACACAGCTGAAAGAAATTAATAAGCAAAATAGCAGAAAGCAAT 540
Db 602 TGTAAATGTGCAACTAAACACAGCTGAAAGAAATTAATAAGCAAAATAGCAGAAAGCAAT 661
QY 541 GAGTGGTTTACTCTTACTTATTAAGCAATAAGCAAAACCAAAAGCACTATACCAAGTGT 600
Db 662 GAATGGTTTACTCTTACTTATTAAGCAATAAGCAAAACCAAAAGCACTATACCAAGTGT 721
QY 601 GTTCTTTTGGAAACATGATGCTTACTTATTTTAACTTAAAGGAAATAGCACTAGT 660
Db 722 GTTCTTTTGGAAACATGATGCTTACTTATTTTAACTTAAAGGAAATAGCACTAGT 781
QY 661 CCACCAAGAGAGCGAGGCTATTTCTTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA 720
Db 782 CCGCCAAAGGAGCGAGGCTATTTCTTAGTACTGCTGGCTGGAAAACTAACTTTTAA 841
QY 721 AAAGAAGGCGAGCGCCATCTAGTGACAACTATACACTGATGACATGCGGCGGCAAGAACG 780
Db 842 AAAGAAGGCGAGCGCCATCTAGTGACAACTATACACTGATGACATGCGGCGGCAAGAACG 901
QY 781 GTTGAACACACAGTAACTACCTGCGCAGGAAACTTAGCGCGGAGAACTCAAACTAAAAAA 840
Db 902 GTTGAACACACAGTAACTACCTGCGCAGGAAACTTAGCGCGGAGAACTCAAACTAAAAAA 961
QY 841 GAAGTTTCTATTAAAACTACCTTAAAGAGCTGGTGTCATAAAAGAGTAACTCACCAGAG 900
Db 962 GAGGTTTCTATTAAAACTACCTTAAAGAGCTAGTGTCATAAGAGAGTAACTCACCAGAA 1021
QY 901 GACTGATGATGATGAGCGGACAGACAGTTCATTTGAAATGATGCTCAACAGGTGGAGAA 960
Db 1022 GACTGATGATGATGAGCGGACAGACAGTTCATTTGAAATGATGCTCAACAGGTGGAGAA 1081
QY 961 AACCTGCTGAAAAATACGCTAGAGATTTGTCACATACTCTAGCCAGAGACCAAAACAGCA 1020
Db 1082 AACCTGCTGAAAAATACGCTAGAGATTTGTCACATACTCTAGCCAGAGACCAAAACAGCA 1141
QY 1021 TTTGACTTAATTTAGAAAAAGCTGAAACCAAGCAAACTAACCAACTTTTCACTGCCTGAC 1080
Db 1142 TTTGACTTAATTTAGAAAAAGCTGAAACCAAGCAAACTAACCAACTTTTCACTGCCTGAC 1201
QY 1081 ACAGAAACCTGAGAAATTTTGTCTTTTCAATGCTGGAACTATGTTAAAGTTGCGCATGCT 1140
Db 1202 ACAAGAAACCTGAGAAATTTTGTCTTTTCAATGCTGGAACTATGTTAAAGTTGCGCATGCT 1261
QY 1141 ATTTGCTGTGTTTTTAAACAGACAGGAGCAAAAGAAATGCTGTTTATTTTCATGAGCA 1200
Db 1262 ATTTGCTGTGTTTTTAAACAGACAGGAGCAAAAGAAATGCTGTTTATTTTCATGAGCA 1321
QY 1201 GCCAGACAGGCAAAATCTATTATTGCAAGCCATAGCACAAGCAGTTCGGCAATGTTGT 1260
Db 1322 GCCAGTACAGGCAAAATCTATTATTGCAAGCCATAGCACAAGCAGTTCGGTAATGTTGT 1381
QY 1261 TGTATATATGAGCCAAATGTTAACTTTTCAATTTATGACTGTATCCAAACAGAACTTGAT 1320
```


Db	1382	TGTTATATGTCAGCAATGTGAATCTTCCATTCAATGACTGTACCAACAAGAACTTAATT	1441
Qy	1321	TGGGTAGAAGAAGCTGGTAACTTTTGGACAGCAAGTAAACAGATTTAAAGCATTTCCTCT	1380
Db	1442	TGGGTGGAAGAAGCTGGTAACTTTTGGACAGCAAGTAAACAGATTTAAAGCATTTCCTCT	1501
Qy	1381	GGTCAAACTATTGCGCATTTGATCAAAAGGAAAGGAGGAGCAAGATTTGAAACCAACCA	1440
Db	1502	GGTCAAACTATTGCGCATTTGATCAAAAGGAAAGGAGGAGCAAGATTTGAAACCAACCA	1561
Qy	1441	GTCAATGATGACCAAAATGAGAACTTACAGTGGTCAAGATAGGCTGCGAAGAAAGACCA	1500
Db	1562	GTCAATGATGACCAAAATGAGAACTTACAGTGGTCAAGATAGGCTGCGAAGAAAGACCA	1621
Qy	1501	GAACACATCAACCAATTCAGAGACAGAACTTTAAACATTCATCAACATACCTTGCCT	1560
Db	1622	GAACACATCAACCAATTCAGAGACAGAACTTTAAACATTCATCAACATACCTTGCCT	1681
Qy	1561	GGTGACTTTGGTTGGTTGA CAATAATGAATGSCCCATGATTTGTGCTTGGTTGAAG	1620
Db	1682	GGTGACTTTGGTTGGTTGA CAATAATGAATGSCCCATGATTTGTGCTTGGTTGAAG	1741
Qy	1621	AATGGTTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTCTGATTGG	1680
Db	1742	AATGGTTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTCTGATTGG	1801
Qy	1681	TCAGAAACTGGGCGGAGCCAAAGTGCCAACTCTTAAATTTACTAGGTTCCGACACGC	1740
Db	1802	ACGAAAACTGGGCGGAGCCAAAGTGCCAACTCTTAAATTTACTAGGTTCCGACACGC	1861
Qy	1741	TCACCAATTACGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTAATCCACTT	1800
Db	1862	TCACCATTTACGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTAATCCACTT	1921
Qy	1801	GCATCGATCTCGAGACCTGGCTTTAGAGCTTTGAGACACCAATATCTCTGTTGG	1860
Db	1922	GCATCGATCTCGAGACCTGGCTTTAGAGCTTTGAGACACCAATATCTCTGTTGG	1981
Qy	1861	GGCACTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGTCAA	1920
Db	1982	GGCACTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGTCAA	2041
Qy	1921	CTGAGCCCAACTGGTTCAGAGATCGAGAGGATTTGAGACCGTCTCGTGCGGAACCG	1980
Db	2042	CTGAGCCCAACTGGTTCAGAGATCGAGAGGATTTGAAAGCGTGTCTCGTGCGGAACCG	2101
Qy	1981	TTCAAGAAAGACTTCAGCGAGCGCTGAACTGCGACTAA	2019
Db	2102	TTGAAGAGAGACTTCAGCGAGCGCTGAACTTGGACTAA	2140
RESULT 13			
HOU34255			
LOCUS	HOU34255		
DEFINITION	Hamster parvovirus DNA.		
ACCESSION	U34255		
VERSION	U34255.1 GI:1464792		
KEYWORDS	.		
SOURCE	Hamster parvovirus		
ORGANISM	Hamster parvovirus		
REFERENCE	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.		
AUTHORS	1 (bases 1 to 4773)		
	Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L.,		
	Franklin,C.L., Hook,R.R. Jr. and Riley,L.K.		
TITLE	Molecular characterization of newly recognized rodent parvoviruses		
JOURNAL	J. Gen. Virol. 77 (Pt 5), 899-911 (1996)		
MEDLINE	96201434		
PUBMED	8609486		
REFERENCE	2 (bases 1 to 4773)		
AUTHORS	Besselsen,D.G.		
	Direct Submission		
TITLE	Submitted (17-AUG-1995) David G. Besselsen, Department of		
JOURNAL	Veterinary Pathology, University of Missouri-Columbia, W213		

FEATURES		Veterinary Medicine Building, Columbia, MO 65211, USA	
Source		Location/Qualifiers	
		1. .4773	
		/organism="Hamster parvovirus"	
		/mol_type="genomic DNA"	
		/specific_hosts="Mesocricetus auratus"	
		/db_xref="taxon:42843"	
ORIGIN			
Query Match		92.2%; Score 1862.2; DB 14; Length 4773;	
Best Local Similarity		95.1%; Pred. No. 0;	
Matches 1921;		Conservative 0; Mismatches 98; Indels 0; Gaps 0;	
Qy	1	ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAAA	60
Db	122	ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAAACCAACCACTGGTTAAAGGAAA	181
Qy	61	AGTAACCAAGAGTGTCTCATTTTAAAAATGAAATGTTCAACTGGAATGAAAA	120
Db	182	AGTAACCAAGAGTGTCTCATTTTAAAAATGAAATGTTCAAGTCAATGGAATA	241
Qy	121	GATATCGGATGAAATAGTTACAAAAAGAGCTGCAGAGGACGAGCTGAAATCTTTACAA	180
Db	242	GATATCGGATGGAATAGTTACAAAAAGAGCTGCAGAGGAGAGCTGAAATCTTTACAA	301
Qy	181	CGAGGAGCGGAAACTTACTTTGGGACCAAGCGAGGACATGGAATGGGAAACCAAGTGGAT	240
Db	302	CGAGGAGCGGAAACTTACTTTGGGACCAAGCGAGGACATGGAATGGGAAATCTTCAGTGGAT	361
Qy	241	GAAATGACCAAAAGCAAGTATTCATTTTGGTTTGGTTTAAATGTTTAAATGTTTAAATGAA	300
Db	362	GAACTAACCAAAAGCAAGTATTCATTTTGGTTTGGTTTAAATGTTTAAATGTTTAAATGAA	421
Qy	301	GTGCTTAAACAAAGAAATATATTTCTGCTGTGATGTTAAATTTGGTTTGTGCAACATGAATGG	360
Db	422	GTGCTGAGTACAAAGAAACATAGCACCTAGTGTACTTGTGTTGTCAGCATGAATGG	481
Qy	361	GGAAAAAGCAAGGCTGGCACTGCCATGTAATAATTTGGAGGAAAGGACTTTAGTCAAGCT	420
Db	482	GGAAAAAGCAAGGCTGGCACTGTCACTGTAATAATTTGGAGGCAAGGACTTTAGCCAAGCT	541
Qy	421	CAAGGAAATGTTGGAGAGGCACTAAATGTTTACTGAGCAGAGATGGTTGTTAAACAGCC	480
Db	542	CAAGGAAATGTTGGAGAGGCACTAAATGTTTACTGAGCAGAGATGGTTGTTAAACAGCC	601
Qy	481	TGTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGAACAT	540
Db	602	TGTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTGAGAGATAGCGGAGAACCA	661
Qy	541	GAGTGGTTTACTCTACTTACTTATAGCATAGCAAAACCAAAAAAGACTATATACCAAGTGT	600
Db	662	GAGTGGTTTACTCTACTTACTTATAGCATAGCAAAACCAAAAAAGACTATATACCAAGTGT	721
Qy	601	GTTCCTTTTGGAAACATGATTTGCTTACTATTTTAACTTAAAGAAATATAGCACTAGT	660
Db	722	GTTCCTTTTGGAAATATAGTTGCTTACTATTTTAACTTAAAGAAATATATGCAACAGT	781
Qy	661	CCACCAAGAGAGCGAGGCTATTTCTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA	720
Db	782	CCACCAAGAGAGCGAGGCTATTTCTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA	841
Qy	721	AAAGAGGCGAGCGCCATCTAGTGACAACTATACCTGATGACATGGCGCCAGAAACG	780
Db	842	AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACCTGATGACATGGCGCCAGAAACG	901
Qy	781	GTGAAACACACAGTAACCACTGGGAGGAACTTAAGCGGGCAGGATTCAACTTAATAAA	840
Db	902	GTGAAACACACAGTAACCACTGGGAGGAACTTAAGCGGGCAGGATTCAACTTAATAAA	961
Qy	841	GAGTTTCTTATTAATAACTACCTTAAAGAGCTGGTGCATATAAAGAGTAACTCACCAGAG	900
Db	962	GAGTTTCTTATTAATAACTACCTTAAAGAGCTGGTGCATATAAAGAGTAACTCACCAGAA	1021

QY 901 GACTGGATGATGTCAGCCAGACAGTTCATTGAAATGATGGCTCAACAGGTGGAGAA 960
DB 1022 GACTGGATGATGTCAGCCAGACAGTTCATTGAAATGATGGCTCAACAGGTGGAGAA 1081
QY 961 AACCTGCTGAAAAATACGCTAGAGATTGTACACTAATCTAGCCAGAACCAAAACAGCA 1020
DB 1082 AACCTGCTGAAAAATACGCTAGAGATTGTACACTAATCTAGCCAGAACCAAAACAGCA 1141
QY 1021 TTTGACTTAAATTTTAGAAAAAGCTGAAACACGAGCAAACTAACCAACTTTTCACTGCCCTGAC 1080
DB 1142 TTTGACTTAAATTTTAGAAAAAGCTGAAACACGAGCAAACTAACCAACTTTTCACTGCCGGAC 1201
QY 1081 ACAAGAACCTGCAGAAATTTTGTCTTTTCATGGCTGGAACATATGTTAAAGTTTGGCATGCT 1140
DB 1202 ACAAGAACCTGCAGAAATTTTGTCTTTTCATGGCTGGAACATATATTAAGTTTGGCATGCT 1261
QY 1141 ATTTGCTGTGTTTAAACAGCAAGAGGAGCAAAAGAAATGCTGTGTTTATTTATTCATGAGCA 1200
DB 1262 ATTTGCTGTGTTTAAACAGCAAGAGGAGCAAAAGAAATGCTGTGTTTATTTATTCATGAGCA 1321
QY 1201 GCCAGCAGGCAAACTCTATTATTGCAAGCCATAGCACAAGCAGTTGGCAATGTTGGT 1260
DB 1322 GCCAGTACAGCAAACTCTATTATTGCAAGCCATAGCAGGAGTTGGTAATGTTGGT 1381
QY 1261 TGCTATAATGAGCCAAATGTAAACTTTCCATTTAATGACTGTACCAACAGAACTTGATT 1320
DB 1382 TGCTATAATGAGCCAAATGTAAACTTTCCATTTAATGACTGTACCAACAGAACTTGATT 1441
QY 1321 TGGGTAGAAGAGCTGTGTAATTTGGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT 1380
DB 1442 TGGGTAGAAGAGCTGTGTAATTTGGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT 1501
QY 1381 GGTCAAACTATTCCCATTTGATCAAAAAGAAAGGAGCAGCAAAACAGATTGAAACCAACCA 1440
DB 1502 GGTCAAACTATTCCCATTTGATCAAAAAGAAAGGAGCAGCAAAACAGATTGAAACCAACCA 1561
QY 1441 GTCATATGACCAAAATGAGAACTTACAGTGTGTCAGAAATAGGCTCGGAAGAAAGACCA 1500
DB 1562 GTCATATGACCAAAATGAGAACTTACAGTGTGTCAGAAATAGGCTCGGAAGAAAGACCA 1621
QY 1501 GAACACACTCAACCAATCAGACAGAGATGTTTAACTTCACTAATCACTAATCACTAATGCTCT 1560
DB 1622 GAACACACTCAACCAATCAGACAGAGATGTTTAACTTCACTAATCACTAATCACTAATGCTCT 1681
QY 1561 GGTGACTTTGTTTGTGTCAGAAATGAATGGCCCATGATTGCTGCTGTTGTTGTTAAAG 1620
DB 1682 GGTGACTTTGTTTGTGTCAGAAATGAATGGCCCATGATTGCTGCTGTTGTTGTTAAAG 1741
QY 1621 AATGGTTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTCTGATTGG 1680
DB 1742 AATGGTTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTCTGATTGG 1801
QY 1681 TCAGAAACTCTGGCGGAGCCAAAGGTGCAATCTCTTAATAATTTACTAGTTTGGCAAGC 1740
DB 1802 ACGAAAACTCTGGCGGAGCCAAAGGTGCGCACTCTTAATAATTTCACTAGTTTGGCAAGC 1861
QY 1741 TCACCAATTCAGCACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTAACTTCCACTT 1800
DB 1862 TCACCAATTCAGCACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTAACTTCCACTT 1921
QY 1801 GCATCGGATCTCGAGGACCTGGCTTTTAGAGCTTTAGAGCCTTGGAGCACCAAACTCTCTGTTGG 1860
DB 1922 GCATCGGATCTCGAGGACCTGGCTTTTAGAGCTTTAGAGCCTTGGAGCACCAAACTCTCTGTTGG 1981
QY 1861 GGCATCTCAGAAACCCAGACACTGGGGAAGCTGGTTTCAAGCTGCGCAAGATGTTCAA 1920
DB 1982 GGCATCTCAGAAACCCAGACACTGGGGAAGCTGGTTTCAAGCTGCGCAAGATGTTCAA 2041
QY 1921 CTGAGCCCAACTTGGTTCAGAGATCGAGGAGATTGTTGAGAGCTGCTTCCGTTGGGAAACCG 1980
DB 2042 CTGAGCCCAACTTGGTTCAGAGATCGAGGAGATTGTTGAGAGCTGCTTCCGTTGGGAAACCG 2101
QY 1981 TTGAAGAAAGACTTTCAGCGAGCCGCTGAACTTGGACTAA 2019

DB 2102 TTGAAGAGAGACTTTCAGCGAGCCGCTGAACTTGGACTAA 2140
RESULT 14
MVU34256
LOCUS
DEFINITION
Mice minute virus DNA.
ACCESSION
U34256
VERSION
U34256.1
GI:1464795
KEYWORDS
Mice minute virus
ORGANISM
Mice minute virus
REFERENCE
1 (bases 1 to 4761)
Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L.,
Franklin,C.L., Hook,R.R. Jr. and Riley,L.K.
Molecular characterization of newly recognized rodent parvoviruses
J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
MEDLINE
96201434
PUBMED
8609486
REFERENCE
2 (bases 1 to 4761)
Besselsen,D.G.
Direct Submission
Submitted (17-AUG-1995) David G. Besselsen, Department of
Veterinary Pathology, University of Missouri-Columbia, W213
Veterinary Medicine Building, Columbia, MO 65211, USA
FEATURES
source
1..4761
/organism="Mice minute virus"
/mol_type="genomic DNA"
/strain="Cutter"
/specific_hosts="Mesocricetus auratus"
/db_xref="taxon:10794"
ORIGIN
Query Match 91.9%; Score 1855.8; DB 14; Length 4761;
Best Local Similarity 94.9%; Pred. No. 0;
Matches 1917; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
QY 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAAA 60
DB 122 ATGGCTGGAAATGCTTACTCTGATGAGTTTGGGAGCAACCACTGGTTAAAGGAAAAA 181
QY 61 AGTAACACGAGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAA 120
DB 182 AGTAACACGAGTGTATCTCATTTGTTTAAAAATGAAGATGTTCAATGAATGAAAA 241
QY 121 GATATCGATGGAATAGTTTACAAAAAGAGCTGCAGGAGCAGAGCTGAAATCTTTACAA 180
DB 242 GATATCGATGGAATAGTTTACAAAAAGAGCTGCAGGAGCAGAGCTAATAATCTTTACAA 301
QY 181 CGAGGAGCGAAACTTACTTTGGGACCAACGAGGACATGGAATGGGAAACCACTGGAT 240
DB 302 CGAGGAGCGAAACTTACTTTGGGACCAACGAGGACATGGAATGGGAACTTTCACTGGAT 361
QY 241 GAAATGACCAAAAAAGCAAGTATTCATTTTGTGTTTAAAAATGTTTATTTGAA 300
DB 362 GAACTAACCAAAAGCAGATTCATTTTGTACTCTTTAGTTTAAAAAGTGTTTATTTGAA 421
QY 301 GTGCTTAAACACAAAGAAATATTTCTGCTGATGTTAAATTTGGTTTGTGCAACATGAATCG 360
DB 422 GTGCTAAGTACAAAGAAACATAGCTCTCTAGTGTATTTAATTTGTTATGTGCGAGCATGATCG 481
QY 361 GGAAAGACCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAAGGACCTTTTACTCAGCT 420
DB 482 GGAAAGACCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAAGGACCTTTTACTAGCCAACT 541
QY 421 CAAGGAAATGTGGAGAGGCAACTAAATGTTTACTGAGCAGATGTTGTTGTAAACAGCC 480
DB 542 CAAGGAAATGTGGAGAGGCACTAAATGTTTACTGAGCAGATGTTGTTGTAAACAGCC 601
QY 481 TGTAAATGTCAACTAACACCACTGAAAGAAATTTAACTAAGAAATAGCAGAAAGCAAT 540

```
Db 602 TGCAGTGTGAGCTATCACAGCGGAAAGAAATTAAGCTGAGAGAAATAGCGGAAGACCA 661
Qy 541 GAGTGGGTACTCTACTTACTTATAAGCATAAGCAACCAAAAGAGACTATACCAAGTGT 600
Db 662 GAATGGGTCACTTGTCTTACTTATAGCATAAGCAACCAAAAGAGACTATACCAATGT 721
Qy 601 GTTCTTTTGGAAACATGATGTCTTACTTATTTTAACTTAAAGAAATAAGCACTAGT 660
Db 722 GTTGTGTTTGGAAATATGATGTCTTACTTATTTTAACTTAAAGAAATAATGCACTAGT 781
Qy 661 CCACCAAGAGCGGAGCTATTTCTTAGCAGTACTCTGGCTGGAAGAACTAATCTTTTA 720
Db 782 CCACCAAGGACGAGGCTATTTCTTAGCAGTACTCTGGCTGGAAGAACTAATCTTTTA 841
Qy 721 AAAGAAGGCGAGGCGCATCTAGTGAGCAAACTATACACTGATGACATCGCGCCAGAAACG 780
Db 842 AAAGAAGGCGAAAGCCATCTAGTGAGCAAACTATACACTGATGACATCGCGCCAGAAACG 901
Qy 781 GTTGAACCAACAGTAAACCACTGCGCAGGAAACTAAGCGCGCAGAAATTCAAAATAAAAA 840
Db 902 GTTGAACCAACAGTAAACCACTGCGCAGGAAACTAAGCGCGCAGAAATTCAAAATAAAAA 961
Qy 841 GAAGTTTCTATTAATACTACACTTAAGAGCTGGTGCATTAAGAGCTTAACCTCACAGAG 900
Db 962 GAGGTTTCTATTAATAACCACTTAAGAGCTGGTGCATTAAGAGCTTAACCTCACAGAA 1021
Qy 901 GACTGGATGATGACGAGCAGACAGTTACATTAAGAAATGATGCTCAACCAAGGTGGAGAA 960
Db 1022 GACTGGATGATGACGAGCAGACAGTTACATTAAGAAATGATGCTCAACCAAGGTGGAGAA 1081
Qy 961 AACCTGCTGAAAAATACGCTAGAGATTGTACACTAACTCTAGCCAGAACCAAAACAGCA 1020
Db 1082 AACCTGCTTAAAAATACGCTAGAGATCTGTACGCTAACTCTAGTAGAACCAAAACAGCC 1141
Qy 1021 TTTGACTTAATTTAGAAAAAGCTGAAACGAGCAAACTAACCACTTTTCACTGCCTGAC 1080
Db 1142 TTTGACTTGAATTTAGAAAAAGCTGAAACGAGCAAACTAACCACTTTTCACTGCCTGAC 1201
Qy 1081 ACAAGAACCTGCAAGATTTTGTGTTTTCATGCTGGAACATATGTTAAAGTTTCCCAATGCT 1140
Db 1202 ACAAGAACCTGTAAGATTTTGTGTTTTCATGCTGGAACATATGTTAAAGTTTCCCAATGCT 1261
Qy 1141 ATTTGCTGTGTTTAAACAGACAGAGGAGGCAAAAGAAATCTGTATTTATTCATGACCA 1200
Db 1262 ATTTGCTGTGTTTAAACAGACAGAGGAGGCAAAAGAAATCTGTATTTATTCATGACCA 1321
Qy 1201 GCAGACACAGGCAAACTATATTGCAACAGCCATAGCAACAGCAGTTGGCAATGTTGTT 1260
Db 1322 GCCAGTACAGGCAAACTCATCATTTGCAACAGCCATAGCAACAGCAGTTGGTAATGTTGTT 1381
Qy 1261 TGCTATTAATGCCCATGTAAACTTTTCCATTTAATGACGTGTACCAACAGAACTTGATT 1320
Db 1382 TGCTATTAATGCAGCAATGTGAATTTTCCATTTCAATGACTGCACCAACAAACCTGATT 1441
Qy 1321 TGGGTAGAAGAGCTGTGTAATCTTTGGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT 1380
Db 1442 TGGGTAGAAGAGCTGTGTAATCTTTGGACAGCAAGTAAACCAATTTAAAGCCATTTGCTCT 1501
Qy 1381 GGTCAAACTATTGCAATGATCAAAAAGGAAAGGAGGCAAGCAACAGATTGAACCAACCA 1440
Db 1502 GGTCAAACTATTGCAATGATCAAAAAGGAAAGGAGGCAAGCAACAGATTGAACCAACCA 1561
Qy 1441 GTCATCATGACCACAATAGAACATTACAGTGTGAGATAGGCTGCGAAGAAAGCA 1500
Db 1562 GTCATCATGACCACAATAGAACATTACAGTGTGAGATAGGCTGCGAAGAAAGCA 1621
Qy 1501 GAACACACTCAACCAATCAGACAGAGATGCTTAACATTTCACTTAACATACACTTGCCT 1560
Db 1622 GAACACACTCAACCAATCAGACAGAGATGCTTAACATTTCACTTAACATACACTTGCCT 1681
Qy 1561 GGTGACTTTGGTTGGTTGCAAAAAATGAATGGCCCAATGATTTGTGCTTGGTTGTAAG 1620
```

```
Db 1682 GGTGACTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGTGCTTGGTTGTTAAG 1741
Qy 1621 AATGTTTACAAATCTACCATGCGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCTGATTGG 1680
Db 1742 AATGTTTACAAATCTACCATGCGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCTGATTGG 1801
Qy 1681 TCAGAAAACCTGGCGGAGCCAAAGGTGCCAACTCCTATATAAATTTACTAGTTTCGGCAGCG 1740
Db 1802 TCAGAAAACCTGGCGGAGCCAAAGGTGCCAACTCCTATATAAATTTACTAGTTTCGGCAGCG 1861
Qy 1741 TCACCAATTCACGACACCCGAAAGTAGTCCTCTCAGCGAGAACTATATGCACTTAATCCACTT 1800
Db 1862 TCACCAATTCACGACACCCGAAAGTAGTCCTCTCAGCGAGAACTATATGCACTTAATCCACTT 1921
Qy 1801 GCATCGGATCTCGAGGACCTGCTTTAGAGCCTTTGAGACCTTTGAGACACACCAAAATCTCTCTTTGG 1860
Db 1922 GCATCGGATCTCGAGGACCTGCTTTAGAGCCTTTGAGACCTTTGAGACACACCAAAATCTCTCTTTGG 1981
Qy 1861 GGCACCTGCAAGAACCCAGAACACTCGGGGAAGCTGTTCCAAAGCCTGCCAAGATGTCAC 1920
Db 1982 GGCACCTGCAAGAACCCAGAACACTCGGGGAAGCTGTTCCAAAGCCTGCCAAGATGTCAC 2041
Qy 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTGGTTCGGGAACCG 1980
Db 2042 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTGGTTCGGGAACCG 2101
Qy 1981 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTTGGACTAA 2019
Db 2102 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTTGGACTAA 2140

RESULT 15
LOCUS CQ786765 5121 bp DNA linear PAT 24-MAR-2004
DEFINITION Sequence 1 from Patent WO2004018689.
ACCESSION CQ786765
VERSION CQ786765.1 GI:45721778
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS 1990, R. and Malerba, M.
TITLE Anti-neoplastic viruses
JOURNAL Patent: WO 2004018689-A 1 04-MAR-2004;
BTG INTERNATIONAL LIMITED (GB)
FEATURES
source
1. .5121
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/noEs="Parvovirus H1\MVM (phH1) - Parvovirus H1 with
promoter P4 and left hairpin from VMV"

ORIGIN
Query Match 87.7%; Score 1771; DB 6; Length 5121;
Best Local Similarity 92.3%; Pred. No. 0;
Matches 1864; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

Qy 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTTAAAGGAAAA 60
Db 282 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTTAAAGGAAAA 341
Qy 61 AGTAACCAAGAGTGTCTCATTTGTTTTTAAAAATGAAATCTTCAACTGAATGAAAA 120
Db 342 AGTAACCAAGAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCACTGAATGAAAA 401
Qy 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTCGAGGAGGACGAGCTGAAATCTTTACAA 180
Db 402 GATATCGGATGGAATAGTTTACAAAAAGAGCTCGAGGAGGACGAGCTGAAATCTTTACAA 461
Qy 181 CGAGGAGCGGAAATCTACTTGGGACCAAGCAGAGCATGGAATGGGAAACCACTGATGAT 240
```

Db 462 CGAGGCGGAACTACTTGGACCAAGCGAGCATGGAATGGGAAACCAACAGTGGAT 521
Qy 241 GAAATGACCAAAAGCAAGTATTCATTTTGGTCTTTGTTTAAATAATGTTTATTGAA 300
Db 522 GAAATGACCAAAAGCAAGTATTCATTTTGGTCTTTGTTTAAATAATGTTTATTGAA 581
Qy 301 GTGCTTAACCAAAAGAAATATATTTCTGGTGATGTTAATTTGGTTTGTGCAACATGAATGG 360
Db 582 GTGCTTAACCAAAAGAAATATATTTCTGGTGATGTTAATTTGGTTTGTGCAACATGAATGG 641
Qy 361 GGAAGAAGCCAGGCTGGCACTGCCATGCTACTAATTCGAGGAAGGACTTTAGTCAAGCT 420
Db 642 GGAAGAAGCCAGGCTGGCACTGCCATGCTACTAATTCGAGGAAGGACTTTAGTCAAGCT 701
Qy 421 CAAGGGAATTTGGTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGTTGGTAAACAGCC 480
Db 702 CAAGGGAATTTGGTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGTTGGTAAACAGCC 761
Qy 481 TGTAAATGTCAACTAACAACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGACAAT 540
Db 762 TGTAAATGTCAACTAACAACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGACAAT 821
Qy 541 GAGTGGGTACTTACTTACTTATTAAGCATTAAGCAACCAAAAGGACTATACCAAGTGT 600
Db 822 GAGTGGGTACTTACTTACTTATTAAGCATTAAGCAACCAAAAGGACTATACCAAGTGT 881
Qy 601 GTTCTTTTGGAAACATGATGTTGCTTACTTATTTTAACTAAAGAAATAAGCACTAGT 660
Db 882 GTTCTTTTGGAAACATGATGTTGCTTACTTATTTTAACTAAAGAAATAAGCACTAGT 941
Qy 661 CCACCAAGAGACGGAGGCTATTTTCTTAGCAGTGAATCTGGCTGGAAACCTAACTTTTAA 720
Db 942 CCACCAAGAGACGGAGGCTATTTTCTTAGCAGTGAATCTGGCTGGAAACCTAACTTTTAA 1001
Qy 721 AAAGAAGCGAGCGCCATCTAGTGAAGCAATATATACCTGATGACATGCGGCCAGAAAG 780
Db 1002 AAAGAAGCGAGCGCCATCTAGTGAAGCAATATATACCTGATGACATGAAACCCAGAAAG 1061
Qy 781 GTTGAACACAGATAACCACTGCGCAGAAACTAAGCGCGCAGAAATTCAACTTAAATAA 840
Db 1062 GTCGAGACCAAGTACCACTGACAGAGAGCTAAGCGCGCAGAAATTCAACTTAGAGAG 1121
Qy 841 GAAGTTTCTATTAAACTATACCTTAAAGAGCTGGTGATATAAGAGTAACTCAACAGAG 900
Db 1122 GAGTCTCGATTAAACCACTCAAGAGTTGGTACATAAAGAGTAACTCAACAGAGAA 1181
Qy 901 GACTGGATGATGACGACAGACAGTTTACATTGAATGATGCTCAACAGGTGGAGAA 960
Db 1182 GACTGGATGATGACGACAGACAGTTTACATTGAATGATGCTCAACAGGTGGAGAA 1241
Qy 961 AACTGTCTGAAATAACGCTAGAGATTTGTACACTAATCTAGCCAGAAACCAAAACAGCA 1020
Db 1242 AACTGTCTTAAATAACACTAGAGATCTGTACACTGACTTAGCAAGAACCAAAACAGCC 1301
Qy 1021 TTGTACTTAATTTTAAAAAGCTGAAACCCAGCAAACTAAACCACTTTTCACTGCGCTGAC 1080
Db 1302 TTTGACTTGAATCTGGAATAAGCTGAAACCCAGCAAACTAGCCAACTTTTCCATGGCTAGC 1361
Qy 1081 ACAAGAAGCTGCAAGATTTTGTCTTTTATGCTGGGAACTATGTTAAAGTTTGGCATGCT 1140
Db 1362 ACCAGAAACCTGTGAATCTTTGTGAGCATGGCTGGAACTATATTAAGTCTGCCATGCC 1421
Qy 1141 ATTTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATGCTGTTTATTTTATGAGACCA 1200
Db 1422 ATCTGTTGTGTGAATAGACAGGAGGCAAAAGGAACTGTGCTCTTTTACAGGACCA 1481
Qy 1201 GCCAGCACAGGCAATCTATTATGCAAGGCCATAGCAACAGCAGTGGCAATGTTGGT 1260
Db 1482 GCCAGCACAGGCAATCTATTATGCAAGGCCATAGCAACAGCAGTGGCAATGTTGGT 1541
Qy 1261 TGCTATAATGCGCAATGTAACCTTTTCAATTTAATGACTGTACCAACAGAACTTGATT 1320
Db 1542 TGTTACAATGTCGCAATGTGAACCTTTTCCATTTTAAATGACTGTACCAACAAAAAATGATT 1601

Qy 1321 TGGGTAGAAGAGCTGTTAACTTTGGACAGCAAGTAAACCCAGTTTAAAGCCATTGCTCT 1380
Db 1602 TGGGTGGAAGAGCTGTTAACTTTGGCCAGCAAGTAAACCAATTCAGAGCTATTGTTCT 1661
Qy 1381 GGTCAAACTATTTCGCAATTGATCAAAAAGGAAAGGAGCAAAACAGATTGAACCAACCA 1440
Db 1662 GGCCAAAACCATACGCAATTGATCAAAAAGGAAAGGAGCAAAACAGATTGAACCAACCA 1721
Qy 1441 GTCATCATGACCAAAATGAGAACATTACAGTGGTCTAGATAGCTGCGAGGAAGACCA 1500
Db 1722 GTTATTATGACCAACCAACGAGAACTTACCGTGGTGTAGATAGCTGTGAGGAAGACCA 1781
Qy 1501 GAAACACACTCAACCAATCGAGACAGAAATGCTTAAACATTTCATCTAACAATACCTTGGCT 1560
Db 1782 GAACACACTCAACCAATCGAGACAGAAATGCTTCAACATTTCACCTGACACGCTACACTACCT 1841
Qy 1561 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGTGCTTGGTTGGTAAAG 1620
Db 1842 GGTGACTTTGGTTGGTTGATAGCAAGTACGCAATGGCCTCTGATCTGTGCTTGGTTGGTGAAG 1901
Qy 1621 AATGGTTTACCAATCTACCATGGCAAGCTACTGTCTTAAATGGGCAAAAGTTCTCTGATTGG 1680
Db 1902 AATGGTTTACCAATCTACCATGGCTTGTACTGTCTTAAATGGGCAAAAGTTCTCTGATTGG 1961
Qy 1681 TCAGAAAACTGGGCGGAGGCAAAAGGTGCCAACTCTCTATAAATTTACTAGGTTTCGGCAGCG 1740
Db 1962 TCAGAGGACTGGGCGGAGGCGGAGCTAGACACTCTCTATAAATTCGCTAGGTTCAATGCGC 2021
Qy 1741 TCACCAATTCACGACACCGAAAAAGTACGCTCTCAGCCAGAACTATGCACTTAACCTCACTT 1800
Db 2022 TCACCAATCTCTGACTCCGAGAGTACGCTCTCAGCCAAAACTACGCTCTTACTCCACTT 2081
Qy 1801 GCATCGATCTCGAGGACCTGGCTTTAGAGCTTTGAGCAGACACCAAAATACTCTCTGTTGG 1860
Db 2082 GCATCGGACCTTCGGGACCTTAGCTCTAGAGCTTTGAGCAGACCAAAATACTCTCTGTTGG 2141
Qy 1861 GGCATCTGCAGAAACCCAGAAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAAGATGGTCAA 1920
Db 2142 GGCATCTGCAGAAACCCAAACACTGGGGAAGCTGGTTCCACAGCCTGCCAAAGGTGCTCAA 2201
Qy 1921 CTGAGCCCAACTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGTTCGGTGGGGAACCG 1980
Db 2202 CGGAGCCCAACTTGGTTCGAGATCGAGGCGGATTTGAGAGCTTGTCTTCAAGTCAAGAACAG 2261
Qy 1981 TTCAGAAAGAGACTTCAGCGAGCGCTGAACTGGACTAA 2019
Db 2262 TTGGAGAGCACTTCAACGAGGAGCTGACCTTGGACTAA 2300

Search completed: January 22, 2005, 21:19:27

Job time : 8547.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2005, 23:46:57 ; Search time 950.5 Seconds
(without alignments)
11150.532 Million cell updates/sec

Title: US-10-069-056-12

Perfect score: 2019

Sequence: 1 atggctggaatgcttactc.....agccgtgaactggactaa 2019

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2019	100.0	2019	5	AAD02803 Parvoviru
2	2017.4	99.9	2019	5	AAD02797 Parvoviru
3	2017.4	99.9	5149	12	ADG39767 Minute vi
4	2015.8	99.8	2019	5	AAD02801 Parvoviru
5	2015.8	99.8	2019	5	AAD02805 Parvoviru
6	2015.8	99.8	2019	5	AAD02799 Parvoviru
7	1040.4	51.5	5049	2	AAT15311 Non-atten
8	1040.4	51.5	5049	2	AAT15312 Attenuate
9	1040.4	51.5	5049	2	AAT88324 Attenuate
10	1038.8	51.5	5049	2	AAT88321 Attenuate
11	1038.8	51.5	5049	2	AAT88320 Canine pa
12	901	44.6	3524	1	AAN40252 Sequence
13	307.4	15.2	374	6	ABQ95626 Tumour su
14	285.8	14.2	421	6	ABQ94779 Tumour su
15	284.8	14.1	423	6	ABQ94724 Tumour su
16	284.8	14.1	464	6	ABQ94719 Tumour su
17	284.8	14.1	473	6	ABQ94716 Tumour su
18	284.8	14.1	486	6	ABQ94707 Tumour su
19	284.8	14.1	491	6	ABQ94777 Tumour su
20	284.2	14.1	420	6	ABQ94787 Tumour su
21	283.8	14.1	343	6	ABQ94737 Tumour su

22	283.8	14.1	343	6	ABQ94739	Abq94739 Tumour su
23	283.8	14.1	403	6	ABQ94782	Abq94782 Tumour su
24	283.2	14.0	497	6	ABQ94718	Abq94718 Tumour su
25	280.4	13.9	340	6	ABQ94732	Abq94732 Tumour su
26	278.2	13.8	342	6	ABQ94765	Abq94765 Tumour su
27	273.8	13.6	420	6	ABQ94781	Abq94781 Tumour su
28	269.6	13.4	424	6	ABQ94740	Abq94740 Tumour su
29	263.4	13.0	339	6	ABQ95625	Abq95625 Tumour su
30	261.6	13.0	424	6	ABQ94778	Abq94778 Tumour su
31	163.8	8.1	468	6	ABQ94873	Abq94873 Tumour su
32	162.2	8.0	451	6	ABQ94868	Abq94868 Tumour su
33	162.2	8.0	457	6	ABQ94867	Abq94867 Tumour su
34	162.2	8.0	465	6	ABQ94874	Abq94874 Tumour su
35	162.2	8.0	469	6	ABQ94869	Abq94869 Tumour su
36	162.2	8.0	472	6	ABQ94880	Abq94880 Tumour su
37	162.2	8.0	515	6	ABQ94793	Abq94793 Tumour su
38	162.2	8.0	516	6	ABQ94807	Abq94807 Tumour su
39	162.2	8.0	516	6	ABQ94794	Abq94794 Tumour su
40	162.2	8.0	516	6	ABQ94800	Abq94800 Tumour su
41	162.2	8.0	517	6	ABQ94811	Abq94811 Tumour su
42	162.2	8.0	530	6	ABQ94791	Abq94791 Tumour su
43	162.2	8.0	530	6	ABQ94799	Abq94799 Tumour su
44	161	8.0	516	6	ABQ94796	Abq94796 Tumour su
45	160.6	8.0	530	6	ABQ94805	Abq94805 Tumour su

ALIGNMENTS

RESULT 1
AAD02803
ID AAD02803 standard; DNA; 2019 BP.
XX
AC AAD02803;

XX
DT 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)

XX
DE Parvovirus non-structure protein 1 (NS1) variant (T394A) DNA.

XX
KW NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy; tumoral disease; gene therapy; mutant; mutein; variant; ds.

XX
OS Parvovirus.
OS Synthetic.

XX
FH Key Location/Qualifiers
FT CDS I..2019

FT mutation /product= "NS1 variant (T394A) protein"
FT /replace(1180, A)
FT /*tag= b

XX
EP1077260-A1.

XX
21-FEB-2001.

XX
13-AUG-1999; 99EP-00115161.

XX
13-AUG-1999; 99EP-00115161.

XX
(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX
Nuesch J, Rommelaere J;

XX
WPI; 2001-212717/22.

XX
P-PSDB; AAY72708.

XX
PT Novel parvovirus non-structure protein variant, useful for treating tumoral diseases, has a shifted equilibrium between DNA replication and transcription activities, and cytotoxic activity.

XX
PS Claim 7; Page 22-24; 41pp; English.

xx The present sequence is a DNA encoding parvovirus non-structure protein 1
CC (NS1) variant (739A). The invention relates to the variants of the
CC parvovirus non-structure protein (NS1) having a shifted equilibrium
CC between the DNA replication and transcription activities, and the
CC cytotoxicity activity. These variants are useful as toxins for treating
CC tumoral diseases. The variant DNAs are useful as vectors for gene
CC therapy. (Updated on 06-AUG-2003 to correct OS field.)
xx
SQ Sequence 2019 BP; 697 A; 389 C; 453 G; 480 T; 0 U; 0 Other;

Query Match 100.0%; Score 2019; DB 5; Length 2019;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2019; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCTGGAATGCTTACTCTGATGAAGTGTGGGAGCAACCACTGGTTAAAGGAAAA 60
Db |||||
Qy 1 ATGGCTGGAATGCTTACTCTGATGAAGTGTGGGAGCAACCACTGGTTAAAGGAAAA 60
Db |||||
Qy 61 AGTAAACGAGGAGTGTCTCTATTTGTTTTTAAAAATGAAATGTCAACTGAATGGAAAA 120
Db |||||
Qy 61 AGTAAACGAGGAGTGTCTCTATTTGTTTTTAAAAATGAAATGTCAACTGAATGGAAAA 120
Db |||||
Qy 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGCAGCTGAAATCTTTACAA 180
Db |||||
Qy 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGGAGCAGCTGAAATCTTTACAA 180
Db |||||
Qy 181 CGAGGAGCGGAACTACTTGGGACCAAGGAGGACATGGAATGGGAAACCAAGTGGAT 240
Db |||||
Qy 181 CGAGGAGCGGAACTACTTGGGACCAAGGAGGACATGGAATGGGAAACCAAGTGGAT 240
Db |||||
Qy 241 GAAATGACCAAAAAAGCAAGTATTCATTTTGTCTTTTAAATAATGTTTATTTGAA 300
Db |||||
Qy 241 GAAATGACCAAAAAAGCAAGTATTCATTTTGTCTTTTAAATAATGTTTATTTGAA 300
Db |||||
Qy 301 GTGCTTAACCAAGAAATATATTCCTGGTGAATGTTAAATGGTTGGTGAACATGAATGG 360
Db |||||
Qy 301 GTGCTTAACCAAGAAATATATTCCTGGTGAATGTTAAATGGTTGGTGAACATGAATGG 360
Db |||||
Qy 361 GGAAGACCAAGCTGGCACTGCATGTAATTAATGGGAAAGGACTTTAGTCAAGCT 420
Db |||||
Qy 361 GGAAGACCAAGCTGGCACTGCATGTAATTAATGGGAAAGGACTTTAGTCAAGCT 420
Db |||||
Qy 421 CAAGGGAATGGTGGAGAGGCACTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCC 480
Db |||||
Qy 421 CAAGGGAATGGTGGAGAGGCACTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCC 480
Db |||||
Qy 481 TGTAAATGCAACTAACACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT 540
Db |||||
Qy 481 TGTAAATGCAACTAACACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT 540
Db |||||
Qy 541 GAGTGGTGTACTCTACTTACTTAAGCATAAGCAACCAAAAGAACTATACCAAGTGT 600
Db |||||
Qy 541 GAGTGGTGTACTCTACTTACTTAAGCATAAGCAACCAAAAGAACTATACCAAGTGT 600
Db |||||
Qy 601 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAAAAAGCACTAGT 660
Db |||||
Qy 601 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAAAAAGCACTAGT 660
Db |||||
Qy 661 CCACCAAGAGACGGAGCTATTTCTTAGCAGTGAATCTGGCTGGAATACTAATTTTAA 720
Db |||||
Qy 661 CCACCAAGAGACGGAGCTATTTCTTAGCAGTGAATCTGGCTGGAATACTAATTTTAA 720
Db |||||
Qy 721 AAAGAGGCGAGCGCACTAGTGAGCAAACTATACACTGATGACATGCGGCGCAGAAACG 780
Db |||||
Qy 721 AAAGAGGCGAGCGCACTAGTGAGCAAACTATACACTGATGACATGCGGCGCAGAAACG 780
Db |||||
Qy 781 GTTGAACCCAGTAACCACTCGGAGGAACTAAGCGCGGAGAAATTCAACTAAAAA 840
Db |||||
Qy 781 GTTGAACCCAGTAACCACTCGGAGGAACTAAGCGCGGAGAAATTCAACTAAAAA 840
Db |||||
Qy 841 GAAGTTTCTATTAATACTACACTTAAGAGCTGGTGCATTAAGAGTAACTCACCAGAG 900
Db |||||

Db 841 GAAGTTTCTATTAATACTACACTTAAGAGCTGGTGCATTAAGAGTAACTCACCAGAG 900
Qy 901 GACTGGATGATGATGAGCAGCAGACAGTTACATTGAATGATGCTCAACAGAGTGGAGAA 960
Db |||||
Qy 901 GACTGGATGATGATGAGCAGCAGACAGTTACATTGAATGATGCTCAACAGAGTGGAGAA 960
Db |||||
Qy 961 AACCTGCTGAAAAAATAGCGCTAGAGATTTGTACACTAAGTCTAGCCAGAACCAAAACAGCA 1020
Db |||||
Qy 961 AACCTGCTGAAAAAATAGCGCTAGAGATTTGTACACTAAGTCTAGCCAGAACCAAAACAGCA 1020
Db |||||
Qy 1021 TTTTGACTTAATTTTACAAAAAGCTGAAACCAAGCAAACTAACCAACTTTTACCTGCCTGAC 1080
Db |||||
Qy 1021 TTTTGACTTAATTTTACAAAAAGCTGAAACCAAGCAAACTAACCAACTTTTACCTGCCTGAC 1080
Db |||||
Qy 1081 ACAAGAACCTGCAGAAATTTTGTCTTTTCATGGCTGGAATCTATTTAAAGTTTCCCATGCT 1140
Db |||||
Qy 1081 ACAAGAACCTGCAGAAATTTTGTCTTTTCATGGCTGGAATCTATTTAAAGTTTCCCATGCT 1140
Db |||||
Qy 1141 ATTTGCTGTGTTTTTAAACAGACAAGGAGGCAAAAGAAATGCTGTTTTTATTTCAAGACCA 1200
Db |||||
Qy 1141 ATTTGCTGTGTTTTTAAACAGACAAGGAGGCAAAAGAAATGCTGTTTTTATTTCAAGACCA 1200
Db |||||
Qy 1201 GCCAGCACAGGCAAAATCTATTATTGCAAGCCATAGCAACAGCAGTGGCAATGTTGCT 1260
Db |||||
Qy 1201 GCCAGCACAGGCAAAATCTATTATTGCAAGCCATAGCAACAGCAGTGGCAATGTTGCT 1260
Db |||||
Qy 1261 TGCTATAATGCAAGCAATGTAAATCTTTCCATTTTAAATGACTGTACCAACAGAACTTTGATT 1320
Db |||||
Qy 1261 TGCTATAATGCAAGCAATGTAAATCTTTCCATTTTAAATGACTGTACCAACAGAACTTTGATT 1320
Db |||||
Qy 1321 TGGGTAGAGAGAGCTGGTAACTTTTGGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT 1380
Db |||||
Qy 1321 TGGGTAGAGAGAGCTGGTAACTTTTGGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT 1380
Db |||||
Qy 1381 GGTCAAACTATTTCGCATTCATCAAAAGGAAAGGAGGAGCAAAAGCAAGTAAACCAACCA 1440
Db |||||
Qy 1381 GGTCAAACTATTTCGCATTCATCAAAAGGAAAGGAGGAGCAAAAGCAAGTAAACCAACCA 1440
Db |||||
Qy 1441 GTCATCATGACCAAAATGAGAACATTACAGTGGTGCAGAAATAGGCTGCGAAGAAAGACCA 1500
Db |||||
Qy 1441 GTCATCATGACCAAAATGAGAACATTACAGTGGTGCAGAAATAGGCTGCGAAGAAAGACCA 1500
Db |||||
Qy 1501 GAACACACTCAACCAATCAGAGACAGAAATGCTTTAACTTAACTTAACTTAACTTAACTTAACT 1560
Db |||||
Qy 1501 GAACACACTCAACCAATCAGAGACAGAAATGCTTTAACTTAACTTAACTTAACTTAACTTAACT 1560
Db |||||
Qy 1561 GGTGACTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGCTTTGGTTGGTAAAG 1620
Db |||||
Qy 1561 GGTGACTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGCTTTGGTTGGTAAAG 1620
Db |||||
Qy 1621 AATGGTTTACCAATCTACCATGCAAGCTACTGTGCTTAAATGGGCGCAAGTTCTGATTGG 1680
Db |||||
Qy 1621 AATGGTTTACCAATCTACCATGCAAGCTACTGTGCTTAAATGGGCGCAAGTTCTGATTGG 1680
Db |||||
Qy 1681 TCAGAAAACTGGGCGAGCCAAAGGTGCAAACTCTCTATTAATTTACTAGTTTGGGACGC 1740
Db |||||
Qy 1681 TCAGAAAACTGGGCGAGCCAAAGGTGCAAACTCTCTATTAATTTACTAGTTTGGGACGC 1740
Db |||||
Qy 1741 TCACCAATTCAGCACACCGGAAAGTACGCTCTCAGCCAGCACTATGCATTAACCTTCCACTT 1800
Db |||||
Qy 1741 TCACCAATTCAGCACACCGGAAAGTACGCTCTCAGCCAGCACTATGCATTAACCTTCCACTT 1800
Db |||||
Qy 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCCTTGGAGCACACCAAAATACTCTGTTGG 1860
Db |||||
Qy 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCCTTGGAGCACACCAAAATACTCTGTTGG 1860
Db |||||
Qy 1861 GGCATCTGCGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCCTGCCAAGTGTCAA 1920
Db |||||
Qy 1861 GGCATCTGCGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCCTGCCAAGTGTCAA 1920
Db |||||
Qy 1921 CTGAGCCCACTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGGTGCTTCGTTGGGAAACCG 1980
Db |||||
Qy 1921 CTGAGCCCACTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGGTGCTTCGTTGGGAAACCG 1980
Db |||||

Qy	1981	TTGAAGAAAGACTTTCAGGAGCCGCTGAACTTGGA	2019
Db	1981	TTGAAGAAAGACTTTCAGGAGCCGCTGAACTTGGA	2019
RESULT 2			
ID	AAD02797 standard; DNA; 2019 BP.		
XX	AAD02797;		
AC	(revised)		
DT	06-AUG-2003 (first entry)		
DT	31-MAY-2001 (first entry)		
XX	Parvovirus non-structure protein 1 (NS1) wild-type DNA.		
DE	NS1: non-structure protein 1; cytostatic; gene therapy; toxin; therapy;		
KW	tumoural disease; gene therapy; ds.		
XX	Parvovirus.		
OS	Key		
FH	Location/Qualifiers		
FT	1..2019		
FT	/*tag= a		
FT	/product= "Parvovirus NS1 protein"		
XX	EPI077260-A1.		
PN	21-FEB-2001.		
XX	13-AUG-1999; 99EP-00115161.		
XX	13-AUG-1999; 99EP-00115161.		
XX	(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.		
XX	Nueesch J, Rommelaere J;		
PI	WPI; 2001-212717/22.		
XX	P-PSDB; AAY72702.		
DR	Novel parvovirus non-structure protein variant, useful for treating		
PT	tumoral diseases, has a shifted equilibrium between DNA replication and		
PT	transcription activities, and cytotoxic activity.		
XX	Disclosure; Fig 1; 41pp; English.		
XX	The present sequence is a wild type DNA encoding parvovirus non-		
CC	structure protein 1 (NS1). The present invention relates to the variants		
CC	of the parvovirus non-structure protein (NS1) having a shifted		
CC	equilibrium between the DNA replication and transcription activities, and		
CC	the cytotoxicity activity. These variants are useful as toxins for		
CC	treating tumoural diseases. The variant DNAs are useful as vectors for		
CC	gene therapy. (Updated on 06-AUG-2003 to correct OS field.)		
XX	Sequence 2019 BP; 698 A; 389 C; 452 G; 480 T; 0 U; 0 Other;		
SQ			
Query Match 99.9%; Score 2017.4; DB 5; Length 2019;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 2019; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	1	ATGGCTGGAATGCTTACTCTGATGAAGTTTGGAGCAACCACTGGTTAAAGGAAAA	60
Db	1	ATGGCTGGAATGCTTACTCTGATGAAGTTTGGAGCAACCACTGGTTAAAGGAAAA	60
Qy	61	AGTAACGAGGAGTGTCTCATTTGTTTTTAAATAAGTAATGTTCAACTGGAATA	120
Db	61	AGTAACGAGGAGTGTCTCATTTGTTTTTAAATAAGTAATGTTCAACTGGAATA	120
Qy	121	GATATCGGATGGAATAGTTTCAAAAAAGAGCTGCAGGAGCAGCTGAAATCTTTACAA	180
Db	121	GATATCGGATGGAATAGTTTCAAAAAAGAGCTGCAGGAGCAGCTGAAATCTTTACAA	180

```
QY 1261 TGCTATAATGCGCAATGTAACCTTTCCATTTAATGACTGTACCAACGAACATTGATT 1320
Db 1261 TGCTATAATGCGCAATGTAACCTTTCCATTTAATGACTGTACCAACGAACATTGATT 1320
QY 1321 TGGGTAGAAGAGCTGGTAACTTTGGACAGCAAGTAAACAGTTTAAAGCATTTCCTCT 1380
Db 1321 TGGGTAGAAGAGCTGGTAACTTTGGACAGCAAGTAAACAGTTTAAAGCATTTCCTCT 1380
QY 1381 GGTCAAACTATTGCGATTGATCAAAAGGAAGGAGGAGCAAGATGAGTGAACCAACCA 1440
Db 1381 GGTCAAACTATTGCGATTGATCAAAAGGAAGGAGGAGCAAGATGAGTGAACCAACCA 1440
QY 1441 GTCATATGACCAAAATGAGAACATTACAGTGGTCAAGATAGCTGCGAAGAAAGACCA 1500
Db 1441 GTCATATGACCAAAATGAGAACATTACAGTGGTCAAGATAGCTGCGAAGAAAGACCA 1500
QY 1501 GAACACACTCAACCAATCAGAGACAGAAATGCTTAACATTCATCAACATACCTTGCCT 1560
Db 1501 GAACACACTCAACCAATCAGAGACAGAAATGCTTAACATTCATCAACATACCTTGCCT 1560
QY 1561 GGTGACTTTGGTTGGTTGACAAAATGAATGCGCCCATGATTTGCTGGTTGGTTAAAG 1620
Db 1561 GGTGACTTTGGTTGGTTGACAAAATGAATGCGCCCATGATTTGCTGGTTGGTTAAAG 1620
QY 1621 AATGGTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCTGATTGG 1680
Db 1621 AATGGTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGGCAAAAGTTCTGATTGG 1680
QY 1681 TCAGAAAACCTGGGCGGAGCAAAAGGTGCCAACTCCTATATAATTTACTAGTTTCGCAAGC 1740
Db 1681 TCAGAAAACCTGGGCGGAGCAAAAGGTGCCAACTCCTATATAATTTACTAGTTTCGCAAGC 1740
QY 1741 TCACCATTCAGCACCCAGCAAGTACGCTCTCAGCCAGCACTATGCACCTAATCCACTT 1800
Db 1741 TCACCATTCAGCACCCAGCAAGTACGCTCTCAGCCAGCACTATGCACCTAATCCACTT 1800
QY 1801 GCATCGGATCTCAGGACCTGGCTTTAGAGCCTTTGAGACACACCAATATCTCTGTTGG 1860
Db 1801 GCATCGGATCTCAGGACCTGGCTTTAGAGCCTTTGAGACACACCAATATCTCTGTTGG 1860
QY 1861 GGCACTGCAGAAACCCAGAACACTGGGGAAAGCTGGTTCCAAAGCCTGCGCAAGATGGTCAA 1920
Db 1861 GGCACTGCAGAAACCCAGAACACTGGGGAAAGCTGGTTCCAAAGCCTGCGCAAGATGGTCAA 1920
QY 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGATTTGAGGCGTCTCGGTGCGGAACCG 1980
Db 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGATTTGAGGCGTCTCGGTGCGGAACCG 1980
QY 1981 TTGAAGAAAGACTTTCAGCGAGCGGCTGAACCTTGGACTAA 2019
Db 1981 TTGAAGAAAGACTTTCAGCGAGCGGCTGAACCTTGGACTAA 2019
```

RESULT 3

```
ID ADG39767
XX ADG39767 standard; DNA; 5149 BP.
AC ADG39767;
XX
XX 11-MAR-2004 (first entry)
DT Minute virus from mouse genomic DNA sequence SEQ ID NO:29.
XX
XX parvovirus; rep; cap; DNA binding domain; capsid interacting domain;
XX recombinant hybrid parvovirus particle;
XX recombinant adeno-associated virus; rAAV; AAV; vaccine; gene therapy;
XX gene; ds.
XX
XX Mice minute virus.
OS
XX
XX WO2003104392-A2.
PN
XX
XX 18-DEC-2003.
PD
```

```
XX 02-DEC-2002; 2002WO-US038423.
PF 18-DEC-2001; 2001US-0341919P.
XX (UYNC-) UNIV NORTH CAROLINA.
PA Samulski RJ, Rabinowitz JE;
XX WPI; 2004-062324/06.
DR New polynucleotides comprising parvovirus rep coding sequences and
PT parvovirus cap coding sequences, useful in producing higher stocks of
PT hybrid parvovirus vectors for delivering therapeutic nucleic acids to a
PT subject.
XX Disclosure; SEQ ID NO 29; 115pp; English.
PS The present invention describes a polynucleotide (I), comprising
CC parvovirus rep coding sequences and parvovirus cap coding sequences. The
CC rep coding sequences encodes a DNA binding domain from a first
CC parvovirus, and a capsid interacting domain from a parvovirus different
CC from the first parvovirus. The cap coding sequence comprises sequences
CC from the different parvovirus. Also described: (1) a vector comprising
CC (1); (2) a cell comprising (1), or parvovirus rep coding sequences and
CC parvovirus cap coding sequences, where the rep coding sequences encode a
CC DNA binding domain from a first parvovirus and a capsid interacting
CC domain from a parvovirus different from the first parvovirus, the cap
CC coding sequences comprise sequences from the different parvovirus, and
CC the rep coding sequences are stably integrated into the genome of the
CC cell; and (3) producing a recombinant hybrid parvovirus particle or adeno
CC -associated virus (rAAV) particle. (I) can be used in vaccines, and in
CC gene therapy. The polynucleotide (I) can be used in producing higher
CC stocks of hybrid parvoviruses or parvovirus vectors, which may be used in
CC the delivery of nucleic acids having biological effect to treat or
CC ameliorate the symptoms associated with any disorder related to gene
CC expression. The polynucleotide may be used to produce a parvovirus vector
CC to express an immunogenic polypeptide in a subject, e.g. for vaccination.
CC The parvovirus vector may also be used to provide an antisense nucleic
CC acid to a cell in vitro or in vivo, or in diagnostic and screening
CC methods. The present sequence is used in the exemplification of the
CC present invention.
XX Sequence 5149 BP; 1718 A; 1045 C; 1124 G; 1262 T; 0 U; 0 Other;
SQ
Query Match 99.9%; Score 2017.4; DB 12; Length 5149;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2018; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA 60
Db 261 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA 320
QY 61 AGTAACGAGAGAGTGTCTCATTTGTTTTTAAATGAAATGTTCAACTGATGGA 120
Db 321 AGTAACGAGAGAGTGTCTCATTTGTTTTTAAATGAAATGTTCAACTGATGGA 380
QY 121 GATATCGGATGGAATAGTTTACAAAAGAGCTCAGAGGAGCAGCTGAAATCTTTACAA 180
Db 381 GATATCGGATGGAATAGTTTACAAAAGAGCTCAGAGGAGCAGCTGAAATCTTTACAA 440
QY 181 CGAGGAGCGGAAACTACTTTGGGACCAAGCGAGGACATGGAATGGGAAACCAAGTGGAT 240
Db 441 CGAGGAGCGGAAACTACTTTGGGACCAAGCGAGGACATGGAATGGGAAACCAAGTGGAT 500
QY 241 GAATGACCAAAAAGCAAGTATTCATTTGTTTCTTGGTTAAATAAATGTTTATTGAA 300
Db 501 GAAATGACCAAAAAGCAAGTATTCATTTGTTTCTTGGTTAAATAAATGTTTATTGAA 560
QY 301 GTGCTTAAACAAAGAAATATATTTCTGTTGATGTTAAATGTTTGTGCAACATGAATGG 360
Db 561 GTGCTTAAACAAAGAAATATATTTCTGTTGATGTTAAATGTTTGTGCAACATGAATGG 620
```

```
QY 361 GGAAGACCAAGCTGGCACTGCCATGTACTAATTGGAGGAAGGACTTTAGTCAAGCT 420
Db 621 GGAAGACCAAGCTGGCACTGGCATGTACTAATTGGAGGAAGGACTTTAGTCAAGCT 680
QY 421 CAAGGGAATGGTGGAGGCAACTAAATGTTTACTTGGAGCAGATGCGTTGGTAAACAGCC 480
Db 681 CAAGGGAATGGTGGAGGCAACTAAATGTTTACTTGGAGCAGATGCGTTGGTAAACAGCC 740
QY 481 TGTAAATGTCAACTAACACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAGCAAT 540
Db 741 TGTAAATGTCAACTAACACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAGCAAT 800
QY 541 GAGTGGTACTTACTTACTTATTAAGCATTAAGCAACCAAAAGACTATACCAAGTGT 600
Db 801 GAGTGGTACTTACTTACTTATTAAGCATTAAGCAACCAAAAGACTATACCAAGTGT 860
QY 601 GTTCTTTTTCGAACATGATTGCTTACTATTTTAACTAAAGAAATTAAGCACTAGT 660
Db 861 GTTCTTTTTCGAACATGATTGCTTACTATTTTAACTAAAGAAATTAAGCACTAGT 920
QY 661 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA 720
Db 921 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA 980
QY 721 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACCTGATGACATGCGGCCAGAAACG 780
Db 981 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACCTGATGACATGCGGCCAGAAACG 1040
QY 781 GTTGAACACACAGTAACCACTGGCAGGAACTTAAGCGCGCAGAAATTCAACTAAAAAA 840
Db 1041 GTTGAACACACAGTAACCACTGGCAGGAACTTAAGCGCGCAGAAATTCAACTAAAAAA 1100
QY 841 GAAATTTCTATTAAAACTACACTTAAGAGCTGGTGCATAAAAAGAGTAACCTCACCAGAG 900
Db 1101 GAAATTTCTATTAAAACTACACTTAAGAGCTGGTGCATAAAAAGAGTAACCTCACCAGAG 1160
QY 901 GACTGGATGATGACGACGACAGCAGTTACATGAAATGATGGCTCAACAGGTGGAGAA 960
Db 1161 GACTGGATGATGACGACGACAGCAGTTACATGAAATGATGGCTCAACAGGTGGAGAA 1220
QY 961 AACTGCTGAAAAATACGCTAGAGATTGTACACTACTCTAGCCAGAACCAAAACAGCA 1020
Db 1221 AACTGCTGAAAAATACGCTAGAGATTGTACACTACTCTAGCCAGAACCAAAACAGCA 1280
QY 1021 TTTGACTTAATTTTAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCCTGAC 1080
Db 1281 TTTGACTTAATTTTAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCCTGAC 1340
QY 1081 ACAAGAACCTGCGAGAAATTTTGTCTTTTCATGGCTGGAACTATGTTAAAGTTTGGCATGCT 1140
Db 1341 ACAAGAACCTGCGAGAAATTTTGTCTTTTCATGGCTGGAACTATGTTAAAGTTTGGCATGCT 1400
QY 1141 ATTGCTGTCTTTTAAACAGACAGGAGGCAAAAGAAATCTGTTTATTTTCATGGACCA 1200
Db 1401 ATTGCTGTCTTTTAAACAGACAGGAGGCAAAAGAAATCTGTTTATTTTCATGGACCA 1460
QY 1201 GCCAGCAGCGCAATCTATTATTGCAAGCCATAGCACAAGCAGTTGGCAATGTTGGT 1260
Db 1461 GCCAGCAGCGCAATCTATTATTGCAAGCCATAGCACAAGCAGTTGGCAATGTTGGT 1520
QY 1261 TGCTATAATGACGCCAATGTAACCTTTCCATTTTAAATGACTGTACCAACAAGAACTTGATT 1320
Db 1521 TGCTATAATGACGCCAATGTAACCTTTCCATTTTAAATGACTGTACCAACAAGAACTTGATT 1580
QY 1321 TGGGTAGAAGAGCTGTAACTTTGGACAGCAAGTAACCAAGTTTAAAGCCATTTGCTCT 1380
Db 1581 TGGGTAGAAGAGCTGTAACTTTGGACAGCAAGTAACCAAGTTTAAAGCCATTTGCTCT 1640
QY 1381 GGTCAAACTATTCGCAATTTGATCAAAAAGGAAAGGAGGAGCAAAACAGATTGAAACCAACCA 1440
Db 1641 GGTCAAACTATTCGCAATTTGATCAAAAAGGAAAGGAGGAGCAAAACAGATTGAAACCAACCA 1700
QY 1441 GTCATCATGACCAAAATGAGAACATTACAGTGGTGCAGAAATAGGCTGCGAAGAAAGACCA 1500
```

```
1701 GTCATCATGACCAAAATGAGAACATTACAGTGGTCAAGATAGGCTGCGAAGAAAGACCA 1760
QY 1501 GAACACACTCAACCAATCAGAGACAGAAATGCTTTAACATTTCATCTAACACATACCTTGCCT 1560
Db 1761 GAACACACTCAACCAATCAGAGACAGAAATGCTTTAACATTTCATCTAACACATACCTTGCCT 1820
QY 1561 GGTGACTTTGGTTGGTTGTCACAAAAATGAATGGCCCATGATTTGTGCTTGGTTGGTAAAG 1620
Db 1821 GGTGACTTTGGTTGGTTGTCACAAAAATGAATGGCCCATGATTTGTGCTTGGTTGGTAAAG 1880
QY 1621 AATGGTTTACCAATCTACCATGCAAGCTACTGTCTTAAATGGGCAAAAGTTCTGATTGG 1680
Db 1881 AATGGTTTACCAATCTACCATGCAAGCTACTGTCTTAAATGGGCAAAAGTTCTGATTGG 1940
QY 1681 TCAGAAAACTGGCGGAGGACCAAGGTGCCAACTCTCTATAAATTTACTAGGTTTCGGCACGC 1740
Db 1941 TCAGAAAACTGGCGGAGGACCAAGGTGCCAACTCTCTATAAATTTACTAGGTTTCGGCACGC 2000
QY 1741 TCACCAATTCAGACACCGAAAAAGTAGCGCTCTCAGCCAGAACTATGCACTAACCTCACTT 1800
Db 2001 TCACCAATTCAGACACCGAAAAAGTAGCGCTCTCAGCCAGAACTATGCACTAACCTCACTT 2060
QY 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGGACACCAAAATCTCTCTGTTCG 1860
Db 2061 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGGACACCAAAATCTCTCTGTTCG 2120
QY 1861 GGCACCTGCAGAAACCCAGAACTCTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920
Db 2121 GGCACCTGCAGAAACCCAGAACTCTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 2180
QY 1921 CTGAGCCCAACTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGTTCGGTTCGGAAACCG 1980
Db 2181 CTGAGCCCAACTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGTTCGGTTCGGAAACCG 2240
QY 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2019
Db 2241 TTGAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA 2279

RESULT 4
AAD02801
ID AAD02801 standard; DNA; 2019 BP.
XX
AC AAD02801;
XX AC
XX 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
XX Parvovirus non-structure protein 1 (NS1) variant (T363A) DNA.
KW NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
tumoural disease; gene therapy; mutant; mutein; variant; ds.
OS Parvovirus.
XX Synthetic.
XX Key
XX CDS
XX Location/Qualifiers
FT 1..2019
FT /tag= a
FT /product= "NS1 variant (T363A) protein"
FT mutation replace(1187, A)
FT /tag= b
XX
XX EP1077260-A1.
XX
XX 21-FEB-2001.
XX
XX 13-AUG-1999; 99BP-00115161.
XX
XX 13-AUG-1999; 99BP-00115161.
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
PA
```


QY 1801 GCATCGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACACCAATACTCTCTGTTCG 1860
Db 1801 GCATCGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACACCAATACTCTCTGTTCG 1860
QY 1861 GGCACTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920
Db 1861 GGCACTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920
QY 1921 CTGAGCCCAACTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGGTGGGAAACCG 1980
Db 1921 CTGAGCCCAACTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGGTGGGAAACCG 1980
QY 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACCTTGAGACTAA 2019
Db 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACCTTGAGACTAA 2019

RESULT 5

AAD02805
ID AAD02805 standard; DNA; 2019 BP.

XX AAD02805;
AC

XX 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)

XX Parvovirus non-structure protein 1 (NS1) variant (T463A) DNA.

DE NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
XX tumoral disease; gene therapy; mutant; mutein; variant; ds.

XX Parvovirus.
OS Synthetic.
OS

FH Key Location/Qualifiers
FT CDS 1..2019

FT /*tag= a
FT /product= "NS1 variant (T463A) protein"
FT mutation replace(1387, A)
FT /*tag= b

XX EPI077260-A1.

XX 21-FEB-2001.

XX 13-AUG-1999; 99EP-00115161.

XX 13-AUG-1999; 99EP-00115161.

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX Nueesch J, Rommelaere J;

XX WPI; 2001-212717/22.

XX P-PSDB; AAY72710.

XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX Claim 7; Page 27-30; 41pp; English.

XX The present sequence is a DNA encoding parvovirus non-structure protein 1
CC (NS1) variant (T463A). The invention relates to the variants of the
CC parvovirus non-structure protein (NS1) having a shifted equilibrium
CC between the DNA replication and transcription activities, and the
CC cytotoxicity activity. These variants are useful as toxins for treating
CC tumoral diseases. The variant DNAs are useful as vectors for gene
CC therapy. (Updated on 06-AUG-2003 to correct OS field.)

XX Sequence 2019 BP; 697 A; 389 C; 453 G; 480 T; 0 U; 0 Other;

XX Query Match 99.8%; Score 2015.8; DB 5; Length 2019;

XX

Best Local Similarity 99.9%; Pred. No. 0; Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCTGGAAATCCTTACTCTGATGAAATTTTGGGAGCAACCAACTGTTTAAAGGAAAA 60
Db 1 ATGGCTGGAAATCCTTACTCTGATGAAATTTTGGGAGCAACCAACTGTTTAAAGGAAAA 60
QY 61 AGTAAACAGGAAGTGTTCATTTTAAAAATGAAATGTTTCAACTGAATGGAATA 120
Db 61 AGTAAACAGGAAGTGTTCATTTTAAAAATGAAATGTTTCAACTGAATGGAATA 120
QY 121 GATATCGGATGGAATAGTTTACAAAAGAGCTCGAGGAGCAGCTGGAATCTTTTCAA 180
Db 121 GATATCGGATGGAATAGTTTACAAAAGAGCTCGAGGAGCAGCTGGAATCTTTTCAA 180
QY 181 CGAGGAGCGGAAACTACTTTGGGACCAAGCGAGGACATGGAATGGGAAACCAAGTGGAT 240
Db 181 CGAGGAGCGGAAACTACTTTGGGACCAAGCGAGGACATGGAATGGGAAACCAAGTGGAT 240
QY 241 GAAATGACCAAAAAGCAAGTATTTTGGTAAATGTTTAAATGTTTAAATGTTTAAATGTTTAA 300
Db 241 GAAATGACCAAAAAGCAAGTATTTTGGTAAATGTTTAAATGTTTAAATGTTTAAATGTTTAA 300
QY 301 GTGCTTAAACACAAAGATATTTTCTGCTGATGTTTAAATGTTTAAATGTTTAAATGTTTAA 360
Db 301 GTGCTTAAACACAAAGATATTTTCTGCTGATGTTTAAATGTTTAAATGTTTAAATGTTTAA 360
QY 361 GGAAGAGCAACGCTGGCACTGCCATGTACTAATTTGGAGGAAAGGACTTTTAGTCAAGCT 420
Db 361 GGAAGAGCAACGCTGGCACTGCCATGTACTAATTTGGAGGAAAGGACTTTTAGTCAAGCT 420
QY 421 CAAGGAAATGTTGGAGAGGCAACTAAATTTTCTGAGAGCAGATGTTTGGTAAACAGCC 480
Db 421 CAAGGAAATGTTGGAGAGGCAACTAAATTTTCTGAGAGCAGATGTTTGGTAAACAGCC 480
QY 481 TGTAAATGTCACACTACACAGCTGAAAGATTTAAACTAGAGAAATAGCAGAGACAAT 540
Db 481 TGTAAATGTCACACTACACAGCTGAAAGATTTAAACTAGAGAAATAGCAGAGACAAT 540
QY 541 GAGTGGTTTACTCTTACTTACTTAAAGCATAAGCAAAACCAAAAGAACTATACCAAGTGT 600
Db 541 GAGTGGTTTACTCTTACTTACTTAAAGCATAAGCAAAACCAAAAGAACTATACCAAGTGT 600
QY 601 GTTCTTTTGGAAACATGATTTGCTTACTTATTTTAACTAAAAAGAAATAAGCACTAGT 660
Db 601 GTTCTTTTGGAAACATGATTTGCTTACTTATTTTAACTAAAAAGAAATAAGCACTAGT 660
QY 661 CCACCAAGAGACGGAGGCTATTTTCTTAGCAGTACTCTGGCTGGGAAATCTAATCTTTTAA 720
Db 661 CCACCAAGAGACGGAGGCTATTTTCTTAGCAGTACTCTGGCTGGGAAATCTAATCTTTTAA 720
QY 721 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG 780
Db 721 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG 780
QY 781 GTTGAACCAACAGTAAACCACTGCGGAGGAACTAAGCGCGGAGAACTTCAAACTAAAAA 840
Db 781 GTTGAACCAACAGTAAACCACTGCGGAGGAACTAAGCGCGGAGAACTTCAAACTAAAAA 840
QY 841 GAAGTTTCTATTAAAACTACCTTAAAGAGCTGGTGCATAAAAAGAGTAACTCACCAGAG 900
Db 841 GAAGTTTCTATTAAAACTACCTTAAAGAGCTGGTGCATAAAAAGAGTAACTCACCAGAG 900
QY 901 GACTGGATGATGATGCGCCAGACAGTTTACATTTGAAATGATGCTCAACAGGTGGAGAA 960
Db 901 GACTGGATGATGATGCGCCAGACAGTTTACATTTGAAATGATGCTCAACAGGTGGAGAA 960
QY 961 AACTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAACCAAAACAGCA 1020
Db 961 AACTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAACCAAAACAGCA 1020
QY 1021 TTTGACTCTAATTTTAGAAAAAGCTGGAACCAAGCAAACTAACCACCTTTTCTGCTGCTGAC 1080
Db 1021 TTTGACTCTAATTTTAGAAAAAGCTGGAACCAAGCAAACTAACCACCTTTTCTGCTGCTGAC 1080

Db 1021 TTTGACTTAATTTTAGAAAAAGCTGAAACCCAGCAAACTAAACCACTTTTCACTGCTGAC 1080
Qy 1081 ACAGAACCTGCAGATTTTGGCTTTTTCATGGCTGGAATATGTTAAAGTTTGGCAATGCT 1140
Db 1081 ACAGAACCTGCAGATTTTGGCTTTTTCATGGCTGGAATATGTTAAAGTTTGGCAATGCT 1140
Qy 1141 ATTGCTGTGTTTTTAAACAGACAGAGGAGGCAAAAGAAATGCTGTTTTTATTTTCAATGGACCA 1200
Db 1141 ATTGCTGTGTTTTTAAACAGACAGAGGAGGCAAAAGAAATGCTGTTTTTATTTTCAATGGACCA 1200
Qy 1201 GCAGACAGGCAAAATCTATTATGTCAGAACCATAGCAGCAAGCAGTGGCAATGTTGTT 1260
Db 1201 GCAGACAGGCAAAATCTATTATGTCAGAACCATAGCAGCAAGCAGTGGCAATGTTGTT 1260
Qy 1261 TGTATTAATGACGCAATGTAATCTTCCATTTTAAATGACTGTACCAACAGAACTTGATT 1320
Db 1261 TGTATTAATGACGCAATGTAATCTTCCATTTTAAATGACTGTACCAACAGAACTTGATT 1320
Qy 1321 TGGGTAGAAAGAGCTGTTAACTTTTGGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT 1380
Db 1321 TGGGTAGAAAGAGCTGTTAACTTTTGGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT 1380
Qy 1381 GGTCAAACTATTGCAATGATCAAAAGAAAGGAGCAAGCAAGATTTGAACCAACCA 1440
Db 1381 GGTCAAGCTATTGCAATGATCAAAAGAAAGGAGCAAGCAAGATTTGAACCAACCA 1440
Qy 1441 GTCATATGACCAATGAGACATTTACAGTGTGTCAGATAGCTGCGAAGAAAGACCA 1500
Db 1441 GTCATATGACCAATGAGACATTTACAGTGTGTCAGATAGCTGCGAAGAAAGACCA 1500
Qy 1501 GAACACACTCAACCAATCAGACAGAGATGTTAACTTCACTTAACACATACCTTGCT 1560
Db 1501 GAACACACTCAACCAATCAGACAGAGATGTTAACTTCACTTAACACATACCTTGCT 1560
Qy 1561 GGTGACTTTGGTTGGTTGACAAATGAATGCGCCATGATTTGCTGCTGTTGTTAAAG 1620
Db 1561 GGTGACTTTGGTTGGTTGACAAATGAATGCGCCATGATTTGCTGCTGTTGTTAAAG 1620
Qy 1621 AATGGTTTACCAATCTACATGCGAAGCTACTGTCTAAATGGGCAAAAGTTCTGATTGG 1680
Db 1621 AATGGTTTACCAATCTACATGCGAAGCTACTGTCTAAATGGGCAAAAGTTCTGATTGG 1680
Qy 1681 TCAGAAACTGGGCGAGCCAAAGGTGCCAACTCCTATAAATTTACTAGTTTGGCAGCG 1740
Db 1681 TCAGAAACTGGGCGAGCCAAAGGTGCCAACTCCTATAAATTTACTAGTTTGGCAGCG 1740
Qy 1741 TCACCATTCAGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800
Db 1741 TCACCATTCAGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800
Qy 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGAGGACACACCAATATCTCTGTTGG 1860
Db 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGAGGACACACCAATATCTCTGTTGG 1860
Qy 1861 GGCACTCGAGAAACCCAGAACACTGGGAAAGCTGGTTCCAAAGCCTCCAAAGATGGTCAA 1920
Db 1861 GGCACTCGAGAAACCCAGAACACTGGGAAAGCTGGTTCCAAAGCCTCCAAAGATGGTCAA 1920
Qy 1921 CTGAGCCCAACTTGGTCAGAGATCGAGAGGATTTGAGAGGCTGCTTCGCTGCGGACCG 1980
Db 1921 CTGAGCCCAACTTGGTCAGAGATCGAGAGGATTTGAGAGGCTGCTTCGCTGCGGACCG 1980
Qy 1981 TTGAAGAAAGACTTTCAGGAGCCGCTGAACCTTGACTAA 2019
Db 1981 TTGAAGAAAGACTTTCAGGAGCCGCTGAACCTTGACTAA 2019

RESULT 6

AAD02799

ID AAD02799 standard; DNA; 2019 BP.

XX

AC

XX

DT 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX Parvovirus non-structure protein 1 (NS1) variant (S283A) DNA.
DE NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoural disease; gene therapy; mutant; mutein; variant; ds.
XX Parvovirus.
OS Synthetic.
XX Location/Qualifiers
FH 1..2019
FT /tag= a
FT /product= "NS1 variant (S283A) protein"
FT mutation
FT replace(847, A)
FT /tag= b
XX EP1077260-A1.
XX 21-FEB-2001.
XX 13-AUG-1999; 99EP-00115161.
XX 13-AUG-1999; 99EP-00115161.
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX Nueesch J, Rommelaere J;
XX WPI; 2001-212717/22.
XX P-PSDB; AAY72704.
XX Novel parvovirus non-structure protein variant, useful for treating
XX tumoral diseases, has a shifted equilibrium between DNA replication and
XX transcription activities, and cytotoxic activity.
XX Claim 7; Page 11-14; 41pp; English.
XX The present sequence is a DNA encoding parvovirus non-structure 1 protein
XX (NS1) variant (S283A). The invention relates to the variants of the
XX parvovirus non-structure protein (NS1) having a shifted equilibrium
XX between the DNA replication and transcription activities, and the
XX cytotoxicity activity. These variants are useful as toxins for treating
XX tumoural diseases. The variant DNAs are useful as vectors for gene
XX therapy. (Updated on 06-AUG-2003 to correct OS field.)
XX Sequence 2019 BP; 698 A; 389 C; 453 G; 479 T; 0 U; 0 Other;
Query Match 99.8%; Score 2015.8; DB 5; Length 2019;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTTAAAGGAAAA 60
Db 1 ATGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTTAAAGGAAAA 60
Qy 61 AGTAACCAAGGAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGGAATA 120
Db 61 AGTAACCAAGGAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGGAATA 120
Qy 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGAGGAGCAGCTGGAATCTTTTACAA 180
Db 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGAGGAGCAGCTGGAATCTTTTACAA 180
Qy 181 CGAGGAGCGGAAATCTTCTGGGACCAAGAGCATGGAATGGGAAACCACTGATGAT 240
Db 181 CGAGGAGCGGAAATCTTCTGGGACCAAGAGCATGGAATGGGAAACCACTGATGAT 240
Qy 241 GAAATGACCAAAAGCAAGATTTCAATTTGATTCTTTGGTTAAAAAATGTTTATTGAA 300
Db 241 GAAATGACCAAAAGCAAGATTTCAATTTGATTCTTTGGTTAAAAAATGTTTATTGAA 300

301 GTGCTTAACACAAAGAAATATATTTCTGCTGATGTTAAATGGTTTGTGCAACATGATGG 360
301 GTGCTTAAACACAAAGAAATATATTTCTGCTGATGTTAAATGGTTTGTGCAACATGATGG 360
361 GGAAGAACACCAAGCTGGCACTGCCATGTACTAAATGGAGCAAGAGACTTTAGTCAAGCT 420
361 GGAAGAACACCAAGCTGGCACTGCCATGTACTAAATGGAGCAAGAGACTTTAGTCAAGCT 420
421 CAAGGGAATATGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCC 480
421 CAAGGGAATATGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCC 480
481 TGTAAATGTGCAACTAACACACAGCTGAAAGAAATTAACATAAGAGAAATAGCAGAGCAAT 540
481 TGTAAATGTGCAACTAACACACAGCTGAAAGAAATTAACATAAGAGAAATAGCAGAGCAAT 540
541 GAGTGGTGTACTCTACTTACTTATAGCATAAGCAACCAAAAGAAAGACTATACCAAGTGT 600
541 GAGTGGTGTACTCTACTTACTTATAGCATAAGCAACCAAAAGAAAGACTATACCAAGTGT 600
601 GTTCTTTTGGAAACATGATTGCTTACTATTTTTTAACTAAAAAGAAATTAAGCACTAGT 660
601 GTTCTTTTGGAAACATGATTGCTTACTATTTTTTAACTAAAAAGAAATTAAGCACTAGT 660
661 CCACCAAGAGACGGAGCTATTTCTTAGCAGTGACTCTGGCTGGAAACTTAATCTTTTA 720
661 CCACCAAGAGACGGAGCTATTTCTTAGCAGTGACTCTGGCTGGAAACTTAATCTTTTA 720
721 AAAGAGGCCAGCGCCATCTAGTGAGCAAACTATACACTGATGACATCGCGCCAGAAACG 780
721 AAAGAGGCCAGCGCCATCTAGTGAGCAAACTATACACTGATGACATCGCGCCAGAAACG 780
781 GTTGAACACCAAGTAAACCACTGGCGAGAACTAAGCGCGCAGAAATTCAAACTAAAAA 840
781 GTTGAACACCAAGTAAACCACTGGCGAGAACTAAGCGCGCAGAAATTCAAACTAAAAA 840
841 GAAGTTTCTAATTAACACTACATTAAGAGCTGGTGATTAAGAGTAACCTCACCAGAG 900
841 GAAGTTTCTAATTAACACTACATTAAGAGCTGGTGATTAAGAGTAACCTCACCAGAG 900
901 GACTGGATGATGACGACGACAGTTCATATTGAATGATGCTCAACACAGGTGGAGAA 960
901 GACTGGATGATGACGACGACAGTTCATATTGAATGATGCTCAACACAGGTGGAGAA 960
961 AACTGCTGAAAAATACGCTAGAGATTGTACACTAATCTAGCCAGAACCAAAACAGCA 1020
961 AACTGCTGAAAAATACGCTAGAGATTGTACACTAATCTAGCCAGAACCAAAACAGCA 1020
1021 TTTGACTTAATTTAGAAAAAGCTGAAACCAAGCAACTAACCACTTTTCACTGCCTGAC 1080
1021 TTTGACTTAATTTAGAAAAAGCTGAAACCAAGCAACTAACCACTTTTCACTGCCTGAC 1080
1081 ACAAGACCTGCAAGATTTTGTCTTTTCATGGCTGGAACTATGTTAAAGTTTCCATGCT 1140
1081 ACAAGACCTGCAAGATTTTGTCTTTTCATGGCTGGAACTATGTTAAAGTTTCCATGCT 1140
1141 ATTTGCTGTGTTTAAACAGACAGGAGCAAAAGAAATGCTGTTTATTTTATTCATGGACCA 1200
1141 ATTTGCTGTGTTTAAACAGACAGGAGCAAAAGAAATGCTGTTTATTTTATTCATGGACCA 1200
1201 GCCAGCAAGGCAAACTATATTGCAAGGCCATAGCAAGCAGTTGGCAATGTTGGT 1260
1201 GCCAGCAAGGCAAACTATATTGCAAGGCCATAGCAAGCAGTTGGCAATGTTGGT 1260
1261 TGCTATATGCGCAATGTAACTTTTCCATTTTAACTGATGACCTACCAAGAACTTGATT 1320
1261 TGCTATATGCGCAATGTAACTTTTCCATTTTAACTGATGACCTACCAAGAACTTGATT 1320
1321 TGGGTAGAAGAGCTGGTAACTTTGGACAGCAAGTAAACAGTTTAAAGCCATTGCTCT 1380
1321 TGGGTAGAAGAGCTGGTAACTTTGGACAGCAAGTAAACAGTTTAAAGCCATTGCTCT 1380
1381 GGTCAAACTATTGCAATTGATCAAAAAAGGAAAGGCGACCAACAGATTGAACCAACACCA 1440

1381 GGTCAAACTATTGCAATTGATCAAAAAGGAAAGGAGGAGCAACAGATTGAACCAACACCA 1440
1441 GTCATCATGACCAACAAATGAGAACTATTACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA 1500
1441 GTCATCATGACCAACAAATGAGAACTATTACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA 1500
1501 GAAACACATCAACCAATCAGAGACAGATGCTTAACTATCATCTAACAACATCTTGCCT 1560
1501 GAAACACATCAACCAATCAGAGACAGATGCTTAACTATCATCTAACAACATCTTGCCT 1560
1561 GGTGACTTTTGGTTTGTGACAAAAATGAATGSCCCATGATTTGTCTTGGTTGGTAAAG 1620
1561 GGTGACTTTTGGTTTGTGACAAAAATGAATGSCCCATGATTTGTCTTGGTTGGTAAAG 1620
1621 AATGTTTACCAATCTACCATGCAAGCTACTGTGCTTAAATGGGCAAAAGTTCTGATTGG 1680
1621 AATGTTTACCAATCTACCATGCAAGCTACTGTGCTTAAATGGGCAAAAGTTCTGATTGG 1680
1681 TCAGAAAACTGGGCGGAGCAAAAGGTGCGCAATCTCTATTAATTTACTAGGTTGCGCACGC 1740
1681 TCAGAAAACTGGGCGGAGCAAAAGGTGCGCAATCTCTATTAATTTACTAGGTTGCGCACGC 1740
1741 TCACCAATTCACGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACCTAACCTCCACTT 1800
1741 TCACCAATTCACGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACCTAACCTCCACTT 1800
1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAAAATACTCTCTGTTGCG 1860
1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAAAATACTCTCTGTTGCG 1860
1861 GGCATGTCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCGCAAGATGGTCAA 1920
1861 GGCATGTCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCGCAAGATGGTCAA 1920
1921 CTGAGCCCACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGCTTCGTTGCGGAACCG 1980
1921 CTGAGCCCACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGCTTCGTTGCGGAACCG 1980
1981 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019
1981 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019

RESULT 7
AAT15311
ID AAT15311 standard; DNA; 5049 BP.
XX AAT15311;
AC AAT15311;
DT 14-OCT-1996 (first entry)
XX
DE Non-attenuated canine parvovirus CPV-39 passage 5 DNA.
XX Parvovirus; dog; vaccine; CPV; myocardial disease; enteric disease; ds;
KW ss.
XX Canine parvovirus.
OS
XX
XX WO9614088-A1.
XX
PD 17-MAY-1996.
XX
XX 02-NOV-1995; 95WO-US014207.
XX
PR 08-NOV-1994; 94US-00336345.
XX
XX (CORR) CORNELL RES FOUND INC.
XX
XX Parriah CR, Gruenberg A, Carmichael LE;
XX WPI; 1996-251556/25.
XX
XX

PT Attenuated CPV strains contg. up to 4 mutation (s) relative to control
PT virus - useful as a veterinary vaccine against CPV disease in animals,
PT such as wild or domestic dogs.
XX Claim 1; Page 21-24; 42pp; English.
XX This viral DNA is isolated from a non- attenuated CPV. The DNA is
CC preferably derived from vB1440. The DNA is cloned into a vector which is
CC used to transfect a host cell. The vector used is preferably pGEM3 or
CC pGEM52. The host cells to be transfected are selected from Norden
CC Laboratory feline kidney cells, mink lung cells, Madin-Darby canine
CC kidney cells or canine A72 cells
XX
XX Sequence 5049 BP; 1791 A; 812 C; 1029 G; 1417 T; 0 U; 0 Other;
Query Match 51.5%; Score 1040.4; DB 2; Length 5049;
Best Local Similarity 71.4%; Pred. No. 4.8e-259;
Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;
Qy 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 60
Db 273 ATGCTGCAACCACTACTGAGGAAGTTATGGAGGAGTAATGTTAAAGAAACAT 332
Qy 61 AGTAAACGAGAGTGTCTCATTTGTTTTTAAATGAAATGTTTCACTGAATGGAAAA 120
Db 333 GCAGAAAAATGAAGCAITTTGTTTTGTTTTTAAATGTGACCAACGTCCTCACTAAATGGAAAG 392
Qy 121 GATATCGGATGGAATAGTTACAAAAAGAGCTGCAGGAGCAGCTGAAATCTTTTACAA 180
Db 393 GATGTTGCTGGAAACAATATACCAACCAATTCAAAATGAAGAGCTTAAATCTTTTAATT 452
Qy 181 CGAGGAGCGGAAACTACTTGGGACCAAGC-----GAGGACATGGAATGGGAACACCA 234
Db 453 AGAGGAGCACAAACAGCAATGATCAACCCGAAGAGAAATGGACTGGGAATCGGAA 512
Qy 235 GTGATGAAATGACAAAAAGCAAGTATTCATTTTTCATTTGTTGTTTAAAAATGTTTA 294
Db 513 GTTGATAGTCTCGCCAAAAAGCAAGTACAACTTTTGATGCAATTAATTAATAAATGCTTT 572
Qy 295 TTTCAAGTGTCTTACACAAAGAAATATTTCTGTCATGTTTAAATGTTTGTGCAACAT 354
Db 573 TTTGAAGTCTTTGTTCTTAAAAATATAGAACCAATGAATGTTGTTGTTTATTTCAACAT 632
Qy 355 GAATGGGAAAAAGACCAAGGCTGGCACTGCCATGCTACTTAATTTGGAGAAAGGACTTTAGT 414
Db 633 GAATGGGAAAAAGATCAAGGCTGSCATTTGTCATGTTTACTTCATAGTAAGAACTTACAA 692
Qy 415 CAAGCTCAAGGAAATGGTGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGGTA 474
Db 693 CAAGCAACTGGTAAATGGCTACGAGACAAATGAATATGTTATGGAGTAGATGGTTGGTG 752
Qy 475 ACAGCTGTAATGTGCAACTAACACCAAGCTGAAGAAATTAACCTTAAGAGAAATAGCAGAA 534
Db 753 ACTCTTTGTCGGTAACTTAACACCACTGAAGAAATTAAGCTCAGAGAAATTTGCAGAA 812
Qy 535 GACAATGAGTGGGTACTCTACTTACTTATAAGCATAGCAACCAAAAGAAAGCTATACC 594
Db 813 GATAGTGAATGGGTGACTATATTAACATACAGACATAGCAACAAAGAAAGACTATGTT 872
Qy 595 AAGTGTGTTCTTTTGGAAACATGATGCTTACTATTTTAACTTAAAGAAAGAAATTAAGC 654
Db 873 AAAATGTTTCATTTTGGAAATATGATAGCATATTTACTTTTAAACAAAGAAAAAATTTGTC 932
Qy 655 ACTAGTCCCAAGAGAGCGAGGCTATTTTCTAGCAGTGAATCTGCTGCTGGAACCTAAC 714
Db 933 CACAT---GACAAAGAAAGTGGCTATTTTAAAGTACTGATCTGTTGGAATTTAAC 989
Qy 715 TTTTAAAGAGCGGAGCGCACTAGTGAGCAAACTATACATGATGACATGCGGCCA 774
Db 990 TTTATGAAGTATCAAGACAGACAGAAATTTGTGACACACTTTTACACTGAACAAATGAACCA 1049
Qy 775 GAAACGGTTGAAACCAAGTAAACCACTGCGGAGAAACTTAAGCGCGGCAGAAATTCAACT 834
|||||

Db 1050 GAAACCGTTGAAACCAACAGTGACGACAGCAAGAAACAAAGCGCGGAGAAATTCAAACT 1109
Qy 835 AAAAAAGAGCTTTCTATTAAACCTACACTTAAAGAGCTGTCATATAAGAGTAACCTCA 894
Db 1110 AAAAGGAAGTGTCAATCAATGATGTTTTCGGGACCTTGGTTAGTAAAGAGTAACATCA 1169
Qy 895 CCAGAGGACTGGATGATGATGACGAGCAGAGTGTACATTGAAATGATGGCTCAACAGGT 954
Db 1170 CTTGAAGACTGGATGATGTTTACAAACAGATAGTTATATTGAAATGATGGCACAACAGGA 1229
Qy 955 GGAGAAAACTGCTGAAAAATAGCTAGAGATTTGTACACTAACTCTAGCAGACAA 1014
Db 1230 GGTGAAAAATCTTTTAAAAAATACACTTGAATTTGTACTTTGACTTTAGCAGAACAAA 1289
Qy 1015 ACAGCATTTGACTTAAATTTTAGAAAAAGCTGAAACGAGCAAACTAACCAACTTTTCACTG 1074
Db 1290 ACAGCATTTGAAATTAATCTTTGAAAAAGCAGATATACTAACTAACTAATCTTTGATCTT 1349
Qy 1075 CTTGACACAGAACTGCGAGAAATTTTGTCTTTTCATGGCTGGAACTATGTTAAAGTTTC 1134
Db 1350 GCAAAATTTCTAGAACATGTCAAAATTTTGTAGAAATGACGATGGAATTTGGAATTAAGTTGT 1409
Qy 1135 CATGCTATTTGCTGTTTAAACAGACAGAGGAGCAAAAGAAATGCTGTTTATTTTCAT 1194
Db 1410 CACGCTATAGCATGTGTTTAAATAGCAAGGTTGTTAAAGAAATACAGTTCTTTTTCAT 1469
Qy 1195 GGAACGAGCAGCAGCAAACTCTATTATTGCAACAGCCATAGCAAGCAGTTGGCAAT 1254
Db 1470 GGAACGAGCAGTACAGGAAATCTATCTGCTCAAGCCATAGCAGCTGTGGTAAAT 1529
Qy 1255 GTTGGTGTCTAATCAGCCAAATGTAATTTTCCATTTTAATGACTGTACCAACAAGAAC 1314
Db 1530 GTTGGTGTCTAATCAGCAAAATGTAATTTTCCATTTTAACTGCTACCAATAAAAT 1589
Qy 1315 TTGATTTGGTAGAAGAGCTGTAATTTTGGACAGCAAGTAAACAGTTTAAAGCCATTT 1374
Db 1590 TTAATTTGGATGGAAGAGCTGTTAACTTTGGTCAACAAAGTTTAACTTAAAGCAATC 1649
Qy 1375 TGCTCTGGTCAAACTATTGTCATTGATCAAAAGAAAGGAGCAAGCAAGATTTGAACCA 1434
Db 1650 TGTTCTGGACAAACAAATTAGATTGATCAAAAGGTAAGGAAGTAAGCAAAATTTGAACCA 1709
Qy 1435 ACACAGTCAATCATGACCAAAATGAGAACATTTACAGTGGTCTAGAAATAGCTCGGAAGAA 1494
Db 1710 ACTCCAGTAAATATGCAACTAATGAAATATAAACAATTTGTGAGAAATTTGGATGGAAGAA 1769
Qy 1495 AGACGAGAACACACTCAACCAATCAGACAGCAAGTCTTAAACATTTCTTAAACATACC 1554
Db 1770 AGACCTGAACATACACAACCAATAGAGACAGAAATGTTGAACATTAAGTTAGTATGTAAG 1829
Qy 1555 TTGCTCTGGTGAATTTGGTTGTAACAAAATGAAATGGCCCATGATTTGTGCTTGGTTG 1614
Db 1830 CTTTCAGGAGACTTTGGTTGGTTGATTAAGAAAGATGSCCTTTAATAATGTCATGGTTA 1889
Qy 1615 GTAAAGATGGTTTCAATCTCATGCGCAAGCTACTGTGCTAAATTTGGGCAAAAGTTCTCT 1674
Db 1890 GTTAAACATGGTTTGAATCAACCATGGCTTAACTATACATCATTTGGGGAAGAAAGTACCA 1949
Qy 1675 GATTGCTCAGAAACTGGGGGAGCCAAAGGTGCCAACTCCTTATAAATTTTACTAGTTGCG 1734
Db 1950 GAATGGATGAAACCTGGGCGGAGCCTTAAATAAAGAAAGGTATAAATTCACCGGTGTC 2009
Qy 1735 GCAGCTCACCATTTCACGACACCGAAAAAGTACGCCCTCTCAGCCAGAACTATATGCACTAACT 1794
Db 2010 AAGACT---TAGAGACACAAAGGCAAGCAATCTCTCAGAGTCAAGCAAGTTCTTAAT 2066
Qy 1795 CCACTTGCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTAGAGCTTTGGAGCACAACCAATATCTCT 1854
Db 2067 CTTCTGACTCCGACCGCTAGTGGACCTTGCACTGGAACCGGTGGAGTACTCCAGATAGCCT 2126
Qy 1855 GTTGGCGGACCTGCAGAAACCCAGAACACTGCGGAGAGCTGTTCCAAAGCCCTGCCAAGAT 1914
Db 2127 ATTGCAGAACTGCAAAATCAACAACTCAAACTTTGGCGTTACTCA-----CAAGAC 2180

Qy 1915 GGTCACGTAGCCCACTGGTCTAGAGATCGAGAGAGATTGAGAGCGTCTCGGTGCG 1974
Db 2181 GTGACGCGAGTCGAGCTGGTCCGAAATAGAGCGACCTGAGAGCCATCTTTACTTCT 2240
Qy 1975 GAACCGTTGAAGAAAGACTTCAGCGA 2000
Db 2241 GAACAAATTGGAAGAAGATTTTCGAGA 2266

RESULT 8

AAT15312
ID AAT15312 standard; DNA; 5049 BP.

XX AC AAT15312;

XX DT 14-OCT-1996 (first entry)

XX DE Attenuated canine parvovirus CPV-39 passage 60 DNA.

XX XX Parvovirus; dog; vaccine; CPV; myocardial disease; enteric disease; ds;

XX KW ss.

XX OS Canine parvovirus.

XX FH Key Location/Qualifiers

FT misc_feature 59

FT FT /*tag= a

FT FT /note= "A, C or T"

FT FT misc_feature 97

FT FT /*tag= b

FT FT /note= "A, G or T"

FT FT misc_feature 4745

FT FT /*tag= c

FT FT /note= "A G or C"

FT FT misc_feature 4881

FT FT /*tag= d

FT FT /note= "A G or T"

XX XX WO9614088-A1.

XX XX 17-MAY-1996.

XX XX 02-NOV-1995; 95WO-US014207.

XX XX 08-NOV-1994; 94US-00336345.

XX XX (CORR) CORNELL RES FOUND INC.

XX XX Parrish CR, Gruenberg A, Carmichael LE;

XX XX WPI; 1996-251556/25.

XX XX Attenuated CPV strains contg. up to 4 mutation (s) relative to control

XX XX virus - useful as a veterinary vaccine against CPV disease in animals,

XX XX such as wild or domestic dogs.

XX XX Claim 2; Page 24-27; 42pp; English.

XX XX This viral DNA is isolated from an attenuated CPV. The DNA is preferably

XX XX derived from vB140. The DNA is cloned into a vector which is used to

XX XX transfect a host cell. The vector used is preferably pGEM3Z or pGEM5Z.

XX XX The host cells to be transfected are selected from Norden Laboratory

XX XX feline kidney cells, mink lung cells, Madin-Darby canine kidney cells or

XX XX canine A72 cells

XX XX Sequence 5049 BP; 1790 A; 814 C; 1029 G; 1416 T; 0 U; 0 Other;

XX XX Query Match 51.5%; Score 1040.4; DB 2; Length 5049;

XX XX Best Local Similarity 71.4%; Pred. No. 4.8e-269;

XX XX Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;

XX XX 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGAGCAACCAACTGGTTAAAGGAAAA 60

Qy

Db 273 ATGCTGGCAACCAAGTATATCTGAGGAAGTTATGGAGGAGTAATGTTTAAAGAAACAT 332
Qy 61 AGTAACCAAGAGAGTGTCTCATTTGTTTTTAAAAATGAAATGTTTCAACTGAATGGAAAA 120
Db 333 GCAGAAAAATGAAGCAATTTTGGTTTTTAAATGTGACAACGTCCTCAACTAAATGGAAAG 392
Qy 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTCGAGAGAGCAGAGCTGAAATCTTTTACAA 180
Db 393 GATGTTGCTGGAAACAACCTATACCAACCAATTCAAAAATGAAGAGCTAACATCTTTAATT 452
Qy 181 CGAGGAGCGGAAACTCTTGGGACCAAGC-----GAGGACATGGAATCGGAAACCA 234
Db 453 AGAGGAGCACAAACAGCAATGGATCAAAACCGAAGAGAGAAATGGACTGGGAATCGGAA 512
Qy 235 GTGGATGAAATGACCAAAAAAGCAAGTATTCATTTTGGTATCTTTTGGTTAAAAAATGTTTA 294
Db 513 GTTGATAGTCTCGCCAAAAAGCAAGTACAACTTTTGGATGCATTAATTTAAAAAATGCTT 572
Qy 295 TTTGAAGTGTCTTAACACAAAGAAATATATTTCTGCTGATGTTTAAATGTTGTTGTCACAT 354
Db 573 TTTGAAGTCTTTGTTTCTAAAAATATAGAACCAAAATGAATGTGTTGTTTATTTCAACAT 632
Qy 355 GAATGGGAAAGCAACCAAGCTGCACTGCATGTACTTAATTCGAGGAAAGGACTTTAGT 414
Db 633 GAATGGGAAAGATCAAGGCTGGCATTTGTCATTTTCTCATAGTAAGAACTTACAA 692
Qy 415 CAAGCTCAAGGGAATGTTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGTTGGTA 474
Db 693 CAAGCAACTGTTAAATGGCTACCGACACAAATGAATATGTTATGGAGTAGATGTTGGTG 752
Qy 475 ACAGCTGTAATGTGCAACTAACACAGCTGAAAGAAATTTAACTAAGAGAAATAGCAGAA 534
Db 753 ACTCTTGTTCGGTAAACTTAACCAACTGAAAGAAATTAAGCTCAGAGAAATTCAGAA 812
Qy 535 GACAATGAGTGGTCTACTCTACTTACTTATAAGCATAGCAACCAACCAAGAAAGTATACC 594
Db 813 GATAGTGAATGGGTGACTATATTAACATACAGACATAAGCAACAAAAAAGAACTATGTT 872
Qy 595 AAGTGTGTTCTTTTGGAAACATGATTCCTTACTATTTTAACTAAAAAGAAATTAAGC 654
Db 873 AATAGTTCATTTTGGAAATATGATAGCATATTTACTTTTAAAGAAATTAAGTTC 932
Qy 655 ACTAGTCCCAAGAGAGCGAGGCTATTTTCTTAGCAGTCACTCTGCTGCGAAACCTAAC 714
Db 933 CACAT---GACAAAAGAAAGTGGCTATTTTAAAGTACTGATTTCTGTTGGAATTTAAC 989
Qy 715 TTTTAAAGAGGCGCGCCCATCTAGTCAGCAAACTATACACTGATGATGATCGCGCCA 774
Db 990 TTTATGAAGTATCAAGACAGACAAATTTGTCAGCACACTTTTACACTGAACAAATGAACCA 1049
Qy 775 GAAACGGTTGAAACCAAGTAAACCACTGCGCAGGAAACTAAGCGCGCAGAAATTCAAACT 834
Db 1050 GAAACCGTTGAAACCAAGTAAACCACTGCGCAGGAAACTAAGCGCGCAGAAATTCAAACT 1109
Qy 835 AAAAAAGAGTGTCTTATTTAAACTACACTTTAAAGAGCTGGTGCATAAAAGAGTAACCTCA 894
Db 1110 AAAAAAGAGTGTCTTATTTAACTCAATCAATGTACTTTGCGGGACTTGGTTAGTAAAGAGTAACATCA 1169
Qy 895 CCAGAGACTGGATGATGTCAGCCAGACAGTTCATTTGAAATGATGGCTCAACAGGT 954
Db 1170 CTTGAAGACTGGATGATGTTTACAAACAGATAGTATATTGAAATGATGGCAACCAACGGA 1229
Qy 955 GGAGAAAACTGCTGAAAAAATACGCTAGAGATTTGTACACTAACTCTTAGCAGAACCAAA 1014
Db 1230 GGTGAAAATCTTTTAAAAAATACACTTGAATTTTGTACTTTGACTTTAGCAGAACAAA 1289
Qy 1015 ACAGCATTTGACTTAAATTTTAGAAAAAGCTGAAACCAAGCAACCACTAACCACTTTTCTG 1074
Db 1290 ACAGCATTTGAAATTAATTAATTTGAAAAAGCAGATAATTAATACTAACTAACTTTGATCTT 1349
Qy 1075 CTTGACACAAAGCACTGACAGAAATTTTGGCTTTTTCATGGCTGGAACTATGTTAAAGTTGC 1134

Db 273 ATGTCTGGCAACCAAGTATCTAGGAAGTTATGAGGGAGTAATAATGGTTTAAAGAAACAT 332
Qy 61 AGTAAACAGGAAGTGTCTCATTTGTTTAAATAAGTAAATGTTCAACTGAATGGAATA 120
Db 333 GCAGAAATCAAGCAATTTTCGTTGTTTAAATGTGACAACTGCAACTTAATGGAAG 392
Qy 121 GATATCGGATGGAATAGTTACAAAAGAGCTGAGAGAGCAAGCTGAAATCTTTACAA 180
Db 393 GATGTTGCGTGGAACTACTATACCAACCAATTCAAAATGAAGAGCTAACTTTAAAT 452
Qy 181 CGAGGAGCGAACTACTCTGGACCAAGC-----GAGGACATGGAATGGGAAACACA 234
Db 453 AGAGGAGCACAACAGCAATGGATCAAAACCGAAGAGAAATGGACTGGGAATCGGAA 512
Qy 235 GTGGATGAAATGACCAAAAAGCAAGTATTCATTTTGTGATTTCTTTGGTTAAAAAATGTTTA 294
Db 513 GTTGATAGTCTCGCCAAAAGCAAGTACAACTTTTGATGCAATTAATTAAGAAATGCTT 572
Qy 295 TTTGAAGTGTCTTAACACAAAGAATATATTTCTGGTGATGTTAAATGGTTTGTGCAACAT 354
Db 573 TTTGAAGTCTTTGTTCTTAAAAATATAGAAACCAATGAATGTGTTGGTTTATTCACAT 632
Qy 355 GAATGGGAAAGACCAAGCTGCACTGCCATGTACTAATTCGAGGAAGGACTTTAGT 414
Db 633 GAATGGGAAAGATCAAGGCTGGCATGTCTATGTTTACTTCATAGTAAGAACTTACAA 692
Qy 415 CAAGCTCAAGGGAAATGGTGGAAAGGCAACTAAATGTTTACTGGAGCAGATGGTTGGTA 474
Db 693 CAAGCAACTGTTAAATGGCTACGCAGACAAATGAATATGTTATGGAGTAGTGGTTGGTG 752
Qy 475 ACAGCCTGTAATGTGCACTAAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAA 534
Db 753 ACTCTTGTTCGGTAAACTTAAACACCACTGAAAGAAATTAAGCTCAGAGAAATTCAGAA 812
Qy 535 GACAATGAGTGGTGTACTCTTACTTATTAAGCATAGCAACCAACCAAAAAGACTATACC 594
Db 813 GATAGTGAATGGGTGACTATATTAACATACAGACATAAGCAACAAACAAAAGACTATGTT 872
Qy 595 AAGTGTGTTCTTTTGGAAACATGATTGCTTACTATTTTAACTTAAAGAAATAAAGC 654
Db 873 AAATGGTTCATTTTGGAAATATGATAGCATATTAATCTTTTAAAGAAATAAATTTGTC 932
Qy 655 ACTAGTCCACAGAGACGAGGCTATTTTCTTAGCAGTGAATCTCTGGCTGGAATACTAAC 714
Db 933 CACAT---GACAAAAGAAAGTGCGTATTTTAAAGTACTGATTCGTTGGAATTTAAC 989
Qy 715 TTTTAAAGAGCGGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCA 774
Db 990 TTTATGAAGTATCAAGACAGACAAATTTGTGAGCACACTTTTACACTGAACAAATGNAACCA 1049
Qy 775 GAAACCGTTGAAACCAAGTAAACACTGCGCAGGAACTAAGCGCGCAGAAATTCAAACT 834
Db 1050 GAAACCGTTGAAACCAAGTAAACACTGCGCAGGAACTAAGCGCGGAGAAATTCAACT 1109
Qy 835 AAAAAAGAGTGTCTATTAATAACTTACATTTAAAGAGCTGGTGATATAAGAGTAACTCA 894
Db 1110 AAAAAAGAGTGTCAATCAATGTACTTTTTCGGGACTTTGGTTAGTAAAGAGTAAACATCA 1169
Qy 895 CCAGAGCACTGGATGATGATGAGCCAGCAGTTCATTTGAAATGATGGCTCAACAGGT 954
Db 1170 CTTGAAGACTGGATGATGTTTACAAACAGATAGTTATATTGAAATGATGGCAGCAACAGGA 1229
Qy 955 GGAGAAAACTGTGAAAAATACGCTAGAGATTTGTACACTACTCTAGCCAGAAACCAA 1014
Db 1230 GGTGAAATCTTTTAAAAAATACACTTGAATTTGTACTTTGACTTTTAGCAGAACAA 1289
Qy 1015 ACAGCACTTTCATTAATTTTAGAAAAAGCTGAAACCAAGCACTAACTTTTCACTG 1074
Db 1290 ACAGCACTTGAATTAATTTAGAAAGCAGATAATACTAACTAACTAACTTTGATCTT 1349
Qy 1075 CTTGACACAAGAACTCTGACGAAATTTTGTCTTTTTCATGGCTGGAACTATGTTAAAGTTTC 1134

Db 1350 GCAAAATCTAGAAACATGTCAAATTTTTAGAAATGCACGGATGGAAATGGATTTAAAGTTGT 1409
Qy 1135 CATGCTATTGCTGTGTTTTTAAACAGACAGAGGCAAAAAGAAATGCTGTTTTATTTCAT 1194
Db 1410 CACGCTATAGCATGTGTTTTTAAATAGCAAGGTGGTAAAGAAATACAGTTCCTTTTTCAT 1469
Qy 1195 GGACCAAGCAGACAGGCAAACTCTATTTNTGCAACAGCCATAGACAGCAGTGGCAAT 1254
Db 1470 GGACCAAGCAGTACAGGAAATCTATCATTTGCTCAAGCCCATAGCAACAGCTGGGTAAT 1529
Qy 1255 GTTGGTGTCTATTAATCAGCCAAATGTAACTTTTCCATTTTAATGACTGTACCAACAAGAC 1314
Db 1530 GTTGGTGTATTAATCAGCAATGTAAATTTTCCATTTTAATGACTGTACCAATTAATAT 1589
Qy 1315 TTGAATTTGGTGAAGAAGCTGGTAACTTTTGGACAGCAAGTAAACAGTTTAAAGCCATT 1374
Db 1590 TTAATTTGGATTGAAGAAGCTGGTAACTTTTGGTCAACAGTTTAACTTTAAAGCAATC 1649
Qy 1375 TGCTCTGGTCAAACTATTGCAATTGATCAAAAAGAAAGGCAAGCAACAGATTTGAACCA 1434
Db 1650 TGTCTCTGGCAAACTTAAGATTTGATCAAAAAGGTAAGAAAGTAAAGCAAAATTTGAACCA 1709
Qy 1435 ACACCAAGTCATCATGACCAACAAATGAGAACATTTACAGTGTGAGATAGGCTCGGAAGAA 1494
Db 1710 ACTCCAGTAAATTAAGCAACTAATGAATAATTAACAATTTGAGAAATTTGAGTGTGAAGAA 1769
Qy 1495 AGACCAAGAACACACTCAACCAATCAGAGACAGAAATGCTTAAACATTTCACTTAACACATACC 1554
Db 1770 AGACCTGAACATACACCAACCAATAGAGACAGAAATGTTGAACATTTAAGTTAGTATGTAAG 1829
Qy 1555 TTGCTCTGGTGAATTTGGTTGGTTGACAAAATTAAGATGGCCCATGATTTGCTGTGGTTG 1614
Db 1830 CTTCCAGGAGACTTTGGTTGGTTGATAAAGAAATGGCCCTTTTAATATGTGATGCTGTTA 1889
Qy 1615 GTAAAGATGGTTTACCAATCTACATGGCAAGCTACTGTGCTAAATGGGCAAGTTCCCT 1674
Db 1890 GTTAAACATGGTTTGAATCAACCATGGCTAACTATACATCATTTGGGGAAGAAATACCA 1949
Qy 1675 GATTGCTCAGAAAACTGGGCGGAGCCAAAGGTGCCAACTCCTATAAAATTTTACTAGTTCG 1734
Db 1950 GAATGGATGAAAACTGGGCGGAGCTTAAATACAAGAAAGGTATAAAATTCACCAGGTTCG 2009
Qy 1735 GCAGCTCACCATTACAGACACGAAAAAGTACGCTCTCAGCCAGAACTATGCACTAACT 1794
Db 2010 AAAGACT---TAGAGACACAAGCGCAAGCAATCTCTCAGAGTCAAGACCAAGTTCTAACT 2066
Qy 1795 CCACTTGGCATCGGATCTCGAGGACCTGGCTTTTAGAGCTTTGGAGCACACCAAAATACTCCT 1854
Db 2067 CCTCTGACTCCGAGCGTAGTGGACCTTGCACTGGAACCGTGGAGTACTCCAGATACGCCT 2126
Qy 1855 GTTGGCGGCACTGCAGAAAAACCCAGAAACACTGGGGAAAGCTGGTTCCAAAGCTCGCCAAGAT 1914
Db 2127 ATTGCAGAACTGCAATCAACAACTCAACCAACTTTGGCGTTACTCA-----CAAAGAC 2180
Qy 1915 GGTCAACTGAGCCCAACTTTGGTTCAGAGATCGAGAGAGATTTGAGAGGCTGCTCGGTGCG 1974
Db 2181 GTGCAAGCAGTCCGAGCTGGTCCGAAATAGAGGCGAGACCTGAGAGCCATCTTTACTTCT 2240
Qy 1975 GAACCGTTGAAGAAAGACTTCAGCGA 2000
Db 2241 GAACAAATGGAAAGAAATTTTCGAGA 2266

RESULT 10
AAT88321
ID AAT88321 standard; DNA; 5049 BP.
XX
AAT88321;
AC
XX
DT 21-MAY-1998 (first entry)
XX
DE Attenuated canine parvovirus genomic DNA.
XX


```
Db 1230 GGTGAAATCTTTAAATAATACACTTGAAATTTGTACTTTGACTTTAGCAAGAACAAA 1289
Qy 1015 ACAGCATTTGACTTAATTTTAGAAAAAGCTGAAACCAAGCAAACTAAACAACTTTTCACTG 1074
Db 1290 ACAGCATTTGAAATTAATTTAGAAAAAGCAGATAATACTAAACTAACTAACTTTGATCTT 1349
Qy 1075 CTTGACACAGAACTTCGAGATTTTTCCTTTTTCATGCTGGAATATGTTAAAGTTTGC 1134
Db 1350 GCAAAATTCAGACATGTCAAAATTTTAGAATGACCGATGGAATTCGATTTAAAGTTTGT 1409
Qy 1135 CATGCTATTTGCTGTTTAAACAGACAGAGGCAAAAGAAATGCTGTTTATTTATTCAT 1194
Db 1410 CAGCTATAGCATGTGTTTAAATAGACAGAGGTGTAAGAAATACAGTCTTTTTCAT 1469
Qy 1195 GGACCAAGCCAGCAGCAAACTTATTTATGCAAGCCATAGCAAGCAGTGTGGCAAT 1254
Db 1470 GGACCAAGCAGTACAGGAAATCTATCTATGCTCAAGCCATAGCAAGCTGTGGTAAT 1529
Qy 1255 GTTGGTTCATATAGCAGCAATGTAACCTTTCATTTTATGACTGTACCAAGCAAC 1314
Db 1530 GTTGGTTCATATAGCAGCAATGTAATTTTCCATTTTATGACTGTACCAATAAAT 1589
Qy 1315 TTGATTTGGGTAGAAAGCTGTTAACTTTTGACAGCAAGTAAACCACTTTAAAGCCATT 1374
Db 1590 TTAATTTGGATGGAAGCTGTTAACTTTTGTGTCACAGTTAATCAATTTAAAGCAATC 1649
Qy 1375 TGCTCTGCTCAAACTATTCGATTTGATCAAAAAGGAAAGCAGCAAAAGATTTGAACCA 1434
Db 1650 TGTTCGACAAACAAATTTAGATTTGATCAAAAAGGTAAGGAATTTGAACCA 1709
Qy 1435 ACACAGTCATATGACACAAATAGAACATTTACAGTGTGTCAGAAATAGGCTGGAAGAA 1494
Db 1710 ACTCCAGTAATATGACAACTAATGAAATTTTAACTTTGTGAGAAATTTGATGTAAGAA 1769
Qy 1495 AGACCAAGACACTCAACCAATCAGACAGAGATGCTTAACTTTCACTTAAACATACC 1554
Db 1770 AGACTGGAATACACAAACCAATAGAGACAGAAATGTTGAACATTTAGTTAGTTAG 1829
Qy 1555 TTGCTGCTGACTTTGTTGTTGTTGACAAAATGAATGGCCCATGATTTGTGCTGTTGTTG 1614
Db 1830 CTTCCAGGAGACTTTGTTGTTGTTGATAAAGAAATGGCTTTTAAATATGTGATGTTA 1889
Qy 1615 GTAAAGAAATGTTACCAATATACATGCGAAGCTACTGTGCTAAATGGGCAAAAGTTCT 1674
Db 1890 GTTAAACATGTTTGAATCAACCATGGCTAACTATACATCATTTGGGGAAGTACCA 1949
Qy 1675 GATTGCTCAGAAACTGGCGGAGCCAAAGTGCCCACTCTCTATAAATTTACTAGTTGCG 1734
Db 1950 GAATGGGATGAAAACTGGCGGAGCCCTAAATAAAGAGGTATAAATTCACCAAGTTGC 2009
Qy 1735 GCAGCTCACCATTTCAGACACACCGAAAGTACGCTCTCAGCCAGAACTATGCATAACT 1794
Db 2010 AAAGACT--TAGAGACACAGCGGCAAGCAATCTCAGAGTCAAGCAAGTTCTTAAT 2066
Qy 1795 CCACTTGATCGGATCTGAGGACCTGGCTTTAGAGCCTTGGAGCACACCAATATCTCT 1854
Db 2067 CTTCTGACTCGGACGTAGTGACCTTGCACTGGAACCGTGGAGTACTCCAGATACGCT 2126
Qy 1855 GTTGGCGGCACTGCAGAAACCCAGAACTGTTGGGGAAGCTGTTCCAAAGCTGCCAAGAT 1914
Db 2127 ATTGCAGAACTGCAAACTCAACAACTCAACAACTTGGCGTTACTCA-----CAAGAC 2180
Qy 1915 GGTCAACTGAGCCCAACTTTGCTCAGAGATCAGAGAGATTTGAGAGCGTGTCTCGGTGCG 1974
Db 2181 GTGCAAGGAGTCCGACGTGCTCGAATAGAGGACGACACCTGAGAGCACTTTACTTCT 2240
Qy 1975 GAACCGTTGAGAAAGACTTCAGGCA 2000
Db 2241 GAACAAATTGGAAGAGATTTTCGAGA 2266
```

RESULT 11
AAT88320

```
ID XX AAT88320 standard; DNA; 5049 BP.
XX AC AAT88320;
XX AC
DT 17-OCT-2003 (revised)
DT 21-MAY-1998 (first entry)
XX Canine parvovirus 39 passage #5 (wild-type).
DE Canine parvovirus; CPV; attenuation; vBI440; vaccine; dog;
XX feline panleukopenia virus; mink enteritis virus; infection; ds.
XX Canine parvovirus; type 2b isolate 39.
XX Key Location/Qualifiers
FH CDS 273..2279
FT /*tag= a
FT /note= "NS1/NS2 coding region"
FT CDS 2286..4541
FT /*tag= b
FT /note= "VP1/VP2 coding region"
XX W09742972-A1.
XX 20-NOV-1997.
XX 06-MAY-1997; 97WO-US007584.
XX 15-MAY-1996; 96US-00647655.
XX (CORR ) CORNELL RES FOUND INC.
XX Parriah CR, Carmichael LE, Gruenberg A;
XX WPI; 1998-008583/01.
XX Canine parvovirus DNA carrying specific attenuating mutation(s) - used as
XX vaccines for protection against parvovirus and feline pan-leukopenia
XX virus infections.
XX Example 8; Page 37-40; 60pp; English.
XX This DNA sequence comprises the genome of virulent canine parvovirus type
XX 2b (CPV-2b) isolate 39, obtained after 5 serial passages in NURK feline
XX kidney host cells. Further passaging has yielded attenuated virus vBI440
XX (ATCC VR 2489) (see AAT88324) from the 60th passage, and a claimed virus
XX (see AAT88321) derived from the 65th passage (ATCC 2528). These
XX respectively contain 4 and 6 mutations in comparison to the virulent 5th
XX passage virus. The DNA from attenuated CPV-2b strains is used for the
XX production of infectious molecular DNA clones, which, in turn, can be
XX transduced into cells to generate master stocks of the virus. The
XX attenuated viruses can be used in dogs as a vaccine to protect against
XX CPV disease, or more generally in cats and minks to protect against
XX feline panleukopenia virus and mink enteritis virus. The vaccines protect
XX against the currently prevalent CPV-2b type (and all extant strains of
XX types 2 and 2a), and provide a long term immune response. (Updated on 17-
XX OCT-2003 to standardise OS field)
SQ Sequence 5049 BP; 1789 A; 813 C; 1030 G; 1417 T; 0 U; 0 Other;

Query Match 51.5%; Score 1038.8; DB 2; Length 5049;
Best Local Similarity 71.3%; Pred. No. 1.3e-268;
Matches 1431; Conservative 0; Mismatches 557; Indels 18; Gaps 4;

Qy 1 ATGCTCGGAATGCTTACTCTGTAAGTTTGGAGCAACCAACTGTTAAAGGAAAA 60
Db 273 ATGCTCGCAACCAAGTATCTAGGAAGTTATGGAGGAGTAAATTTGTTAAGAACAT 332
Qy 61 AGTAACCAAGGAAGTGTCTCATTTGTTTAAATAATGAAATGTTCAACTGAATGGA 120
Db 333 GCAGAAATGACGATTTTCGTTGTTTAAATGTCACACGTCACCACTAAATGGAAG 392
Qy 121 GATATCGGATGGAATAGTTTCAAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTTACA 180
```



```
FT      /*tag= a  
FT      /note= "see AAP40306"  
FT      2107..3522  
FT      /*tag= b  
FT      /note= "see AAP40675"  
XX  
PN      W08402847-A.  
XX  
XX      02-AUG-1984.  
XX  
XX      19-JAN-1984; 84WO-US000063.  
XX  
XX      19-JAN-1983; 83US-00459203.  
PR      06-JAN-1984; 84US-00567968.  
XX  
XX      (AMGE-) AMGEN.  
XX  
XX      FOX GM;  
XX  
DR      WPI; 1984-201354/32.  
XX      P-PSDB; AAP40306, AAP40675.  
XX  
PT      Polypeptide obtd. by recombinant DNA methods - for vaccination against  
PT      parvovirus infections in man and animals.  
XX  
XX      Claim 10; Table II, Page 33-49; 80pp; English.  
XX  
CC      The inventors claim an immunologically active polypeptide for the  
CC      development of vaccinal immunity against parvovirus infection. Also  
CC      claimed are DNA sequences wholly or partly duplicative of defined  
CC      sequences. The polypeptides are used in vaccines for conferring  
CC      protection against parvovirus infections in man and animals. (Updated on  
CC      24-OCT-2003 to standardise OS field)  
XX  
XX      Sequence 3524 BP; 1400 A; 719 C; 648 G; 757 T; 0 U; 0 Other;  
SQ  
  
Query Match      44.6%; Score 901; DB 1; Length 3524;  
Best Local Similarity 69.7%; Pred. No. 1.4e-231;  
Matches 1296; Conservative 0; Mismatches 545; Indels 18; Gaps 5;  
  
QY      4 GCTGGAAATCCTTACTCTGATGAAGTTTTCGGAGACCACTGGTTTAAAGGAAAAGT 63  
DB      13 GCGGAAACACTTACTCGGAAGAGGTACTAAAGCTACCACTGGCTTCAAGATAATGCT 72  
QY      64 AACGAGGAAGTGTCTCATTTGTTTTTAAATAAGTAAATGTTCAACTGAATGGAAGAAT 123  
DB      73 CAAAAGAAGCAATCTCTTATGATTTTAAACACAAAAGTCAATCTAAATGGAAGAAGAA 132  
QY      124 ATCGGATGGAATAGTTACAAAAGAGCTGCAGAGACGAGCTGAAATCTTTACACGA 183  
DB      133 ATTGCTTGGAAATACTACACAAAGATACACAGATCGGAAATGTATAAACCTACAAAGA 192  
QY      184 GGAGCGAAACTCTTGGGACCAAGCGAGCATGGAATGGGAAACCAAGTGGATGAA 243  
DB      193 GGAGCAGAAACATCATGGGACCAAGGCAACAGACATGGAATGGGAATCAGAAATCGACGC 252  
QY      244 ATGACCAAAAAGCAAGTATTCATTTTGTGTTTAAATAAGTATTTTATTTGAAGTG 303  
DB      253 CTCACAAAAGCGCAAGTACTGATTTTGTCTCTTGTGTTAAATAAGTCTCTTTGAAGGT 312  
QY      304 CTTAACAACAAGATAATATTTCTGGTGATGTTAAATTTGTTTGTGCAACATGAATGGGA 363  
DB      313 ATATTGCAAAAGAACCTTAAGTCCAAGTGACTGTCTTCTTACAGCATGAACATGTT 372  
QY      364 AAGACCAAGCTGGGACTGCCATGTACTAATTTGGAGGAAGGACTTTAGTCAAGTCAA 423  
DB      373 CAAGATACTGGCTATCACTGCCATGTACTAGTGGGAAAGGCTTTACAAACGAAGCATG 432  
QY      424 GGGAAATGGTGGAGGCAAGCAACTAAATGTTTACTTGGAGCAGATGTTGGTAAACAGCCTGT 483  
DB      433 GGAATAATGGTTACGAAACAAATTAACAAATTTATGGAGTAGATGTTGAAATATCAATGC 492  
QY      484 AATGTGCAACTAACACCGAGCTGAAAGAAATTTAAACTAAGAGAAATPAGCAGAGAACAATGAG 543  
  
Db      493 AAAGTACCTCTAACACCAAGTTGAAAGAAATAAATTAAGGGAATTTAGCAGAGATGGTGAG 552  
QY      544 TGGGTTACTCTACTTACTTATAAGCATAAAGCAAAACCAAAAAAGACTATACCAAGTGTGT 603  
Db      553 TGGGTATCGCTACTAACTACACTCAAAACAAACATAAAAAACAATATACAAAAATGACT 612  
QY      604 CTTTTTGGAAACATGATGCTTACTATTTTAACTTAAAGAAAGAAATAAGCAGCTAGTCCA 663  
Db      613 CATTTTGGAAATATGATGCTTACTTCTTAATAAAAAAGAAAGAAAGCAACT- ----- 666  
QY      664 CCAAGAGACGGAGCTATTTTCTTAGCAGTACTCTGGCTGGAAAACTAACTTTTTTAAAA 723  
Db      667 GAAAGAGAGCATGGATATATCTCAGCTCAGATTTCTGGCTTCATGACAAATTTCTTAAAA 726  
QY      724 GAAAGGAGCGCCATCTTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACGGTT 783  
Db      727 GAAGCGAGAGACACTTAGTCAGTCACCTATTTTACTGGAAGCAATAAACCTGAAACTGTG 786  
QY      784 GAAACCAAGTAACCACTGCGCAGGAAACTAAGCGCGGAGAGATTTCAAATTAATAAGAA 843  
Db      787 GAAACCAACGGTTACTCAGCTCAGGAAGTTTCCCGAGGCGAGATATACAAACAAAAAGAA 846  
QY      844 GTTCTTATTAAGACTACACTTAAAGAGCTGCTGCATAAAGAGTAACTCTCACCAGAGGAC 903  
Db      847 GTAAGCATAAAAATGCACAAATAGAGACTTGGTTAATAAAGATGTACTAGCATAGAAGGC 906  
QY      904 TGGATGATGATGCGCAGCAGACAGTTACATTTGAAATGATGCTCAACCCAGGTGGAGAAAAAC 963  
Db      907 TGGATGATGACAGATCCAGACAGTTATATAGAAATGATGCTCAAACCGGAGGAGAAAT 966  
QY      964 CTCTGAAAAAATAGCTAGAGATTTTGTACACTAATCTAGCCAGAACCAAAACAGCATTT 1023  
Db      967 TTAATCAAAAATACACTAGAAATAACAACTCTTACTCTAGCAAGAAACAAAAACAGCATAT 1026  
QY      1024 GACTTAATTTTAGAAAAGCTGAAACCAAGCAAACTAACCACTTTTCACTGCCTGACACA 1083  
Db      1027 GACTTAATACTTGAAGGCAAAACCAAGCATGCTACCAACATTTAATAATTTAGCAATACA 1086  
QY      1084 AGAACCTGCAAGATTTTGTCTTTCATGGCTGGAACATATGTTTAAAGTTTGGCATGCTATT 1143  
Db      1087 AGAACATGTAATATTTTCAGATGCACAATTTGNACTACATTTAAAGTCTGCCATGCTATA 1146  
QY      1144 TGCTGTGT-TTTTAAACAGACAAAGAGGCAAAAGA-----AATGCTGTTTATTTTCATGGA 1197  
Db      1147 ACTTGTGCTACTGAAACAGACAAAGGAGGAAAAAGAAATACAAATTCATTTCTATGTCATGGG 1206  
QY      1198 CCAGCCAGCAGCAGCAATCTATTTTGTCAAGCCATAGCAGACAGTTCGCAATGTT 1257  
Db      1207 CCAGCATCAACAGGAAAAAGTATAATTTGCTCAACACATTTGCAAACTTAGTTGGTAATGTT 1266  
QY      1258 GGTGCTTATTAATGAGCAATGTAAACCTTTTCCATTTTAATGACTGTACCAACAGAACTTG 1317  
Db      1267 GGTGCTCAATGAGCCATGTGAACCTTTCCATTTTANTGACTGTACAAATTAATAAATCTTA 1326  
QY      1318 ATTTGGGTAGAAGAAGCTGGTAACTTTTGGACAGCAAGTAAACCAAGTTTAAAGCCATTGC 1377  
Db      1327 ATATGGATTGAAGAAGCAGGAAACTTCTCTAAACCAAGTAAACCAATTTCAAAGCCATATGT 1386  
QY      1378 TCTGGTCAAACTATTCGCAATTGATCAAAAAGGAAAGGCAAGCAACAGATTTGAACCAACA 1437  
Db      1387 TCAGGTCAAACTATTAAGATTGACCAAAAAGGTAAGGAAAGCAAACTTTGAACCAACT 1446  
QY      1438 CCAGTCATCATGATCAACCAAAATGAAACATTTACAGTGGTTCAGAAATAGGCTGCGAAGAAAGA 1497  
Db      1447 CTTGTAATAATGACTCAAAATGAAGACATATCTTAAGTTAGATAGATGGAGGAGAAAGA 1506  
QY      1498 CCAGAACACACTCAACCAATCAGAGACAGAAATGCTTAAACATTTCTATCAACACATACCTTG 1557  
Db      1507 CCAGAACATACACCAACCAATAAGAGACAGAAATGTTTAAACATACACCTTAAACAGAAAACTG 1566  
QY      1558 CTTGGTGACTTTGGTTGGTTTGACAAAAATGAATGGCCCATGATTTGTGCTTGGTGGTA 1617
```

Db 1567 CAGGTGATTTTGGACTTTTAGAAGAACTGTAATGCCACTAATATGTGCTTGGTTGGTA 1626
 Qy 1618 AAGAATGGTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGGCAAGTTCTGTAT 1677
 Db 1627 AAGAAAGGTTACCAAGCAACAATGGCTAGCTATATGTCATCATTTGGGGAATGTACTGTAT 1686
 Qy 1678 TGGTCAGAAACTGGGGAGGCAAGGTGCGCAACTCTTATATAATTTACTAGTTTCGGCA 1737
 Db 1687 TGGTCAGAAAAATTTGGAGGCGCAAAATGTCATTTCCCAATAATACACCA---ACAGAC 1743
 Qy 1738 CGCTCACCAATTCACGACACCGAAGATGACGCTCTCAGCAGAACTATGCACTAATCCA 1797
 Db 1744 TCTCAGATTTCCATCATGAGTGAACCTTGCCAGCGGACATCACTACGACGCAACTCCA 1803
 Qy 1798 CTTGCAATCGATCTCGA---GGACCTGGCTTTAGAGCTTGGAGCACCAACCAATCTCC 1853
 Db 1804 ATACAGGAGGACCTGGATTTAGCTTTAGCCTTGGAGCGGTGGAGCGGCAACACACC 1862

RESULT 13
 ABQ95626
 ID ABQ95626 standard; DNA; 374 BP.
 AC ABQ95626;
 DT 28-OCT-2002 (first entry)
 XX Tumour suppression-related oligonucleotide #1277.
 XX Tumour; cytostatic; antiviral; neuroprotective; nootropic; neuroleptic;
 XX tumour suppression; tumour reversion; apoptosis; viral resistance; human;
 XX viral infection; cell degeneration disease; neurodegeneration; ds;
 XX Alzheimer's disease; schizophrenia; immune disease; inflammatory disease.

OS Homo sapiens.
 XX FR2819824-A1.
 XX 26-JUL-2002.
 XX 23-JAN-2001; 2001FR-00000899.
 XX 23-JAN-2001; 2001FR-00000899.
 XX (MOLE-) MOLECULAR ENGINES LAB SA.
 XX Telerman A, Amson R, Tuijnder M, Susini L;
 XX WPI; 2002-610803/66.
 XX New nucleic acid implicated e.g. in tumor suppression, useful for
 XX diagnosis of tumors, viral infection and cellular degeneration and for
 XX drug screening.

XX Claim 1; Page 360; 623pp; French.
 XX The present invention relates to novel human nucleic acid sequences (I).
 XX The present sequence is one such nucleic acid sequence. Expression of (I)
 XX are implicated in tumour suppression or reversion and apoptosis and viral
 XX resistance. (I) are useful as probes or primers for detecting,
 XX identifying, measuring and/or amplifying nucleic acid sequences, as
 XX antisense reagents and for recombinant production of polypeptides. (I),
 XX polypeptides (II) encoded by (I), vector containing (I), cells containing
 XX these vectors and antibodies (Ab) against (II) are all useful for
 XX treatment/prevention of viral, tumour and cell degeneration diseases
 XX (especially neurodegeneration, such as Alzheimer's disease and
 XX schizophrenia). Analysing the expression of (I) is also useful for
 XX diagnosis and/or prognosis of such diseases. Transgenic animals carrying
 XX (I) are used for studying the aetiology of these diseases (also immune
 XX and inflammatory diseases). Note: In the present specification, SEQ ID 1
 XX to 2280 are claimed in Claim 1, however only SEQ ID 1 to 2270 are shown
 XX in the specification

SQ Sequence 374 BP; 116 A; 57 C; 108 G; 84 T; 0 U; 9 Other;
 Query Match 15.2%; Score 307.4; DB 6; Length 374;
 Best Local Similarity 88.2%; Pred. No. 3.2e-72;
 Matches 329; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
 Qy 174 TTTCACAGAGGAGCGGAAACTTCTGGACCAAAAGCGAGGACATGTAATGGAAACAC 233
 Db 1 TTTCACAGAGGAGCGGAGACCACTTGGACCAAAAGCGAGGACATGTAATGGAGCGC 60
 Qy 234 AGTGAATGAATGACCAAAAGCGAGGACATGTAATTTTGGTTTAAATAATGTTT 293
 Db 61 AGTGAATGACATGACCAAAAGCGAGGACATGTAATTTTGGTTTAAATAATGTTT 120
 Qy 294 ATTTCAGAGTCTTAACACAAAGAAATATTTCTCTGTGTATTTAAATTTGTTGCAACA 353
 Db 121 GTTTGAGTGTCTCAGCACAAAGACATAGCTCTCTAGTATGTTACTTGGTTCTGCAGCA 180
 Qy 354 TGAATGGGAAAAGACCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAGGACCTTAG 413
 Db 181 TGAATGGGAAAAGACCAAGGCTGGCACTGTCTGTGTGTGATTTGGAGGCAAGACTTTAG 240
 Qy 414 TCAAGCTCAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTGTGAGCAGATGTTGT 473
 Db 241 TCAAGCTCAAGGAAATGTTGGAGAGGCACTAAATGTTGTAAGTANATGTTGTG 300
 Qy 474 AACAGCTGTAAATGTGCAACTAACACCACTGAAGAATTAACCTAAGAGAAATAGCAGA 533
 Db 301 GACTGCTGNAATGTNCAACTAACACCACTGAAGAATTAACCTGNNAGAAATAGCAGA 360
 Qy 534 AGACAATGAGTGG 546
 Db 361 GGACAGTGANNNG 373

RESULT 14
 ABQ94779
 ID ABQ94779 standard; DNA; 421 BP.
 XX ABQ94779;
 XX 28-OCT-2002 (first entry)
 XX Tumour suppression-related oligonucleotide #430.
 XX Tumour; cytostatic; antiviral; neuroprotective; nootropic; neuroleptic;
 XX tumour suppression; tumour reversion; apoptosis; viral resistance; human;
 XX viral infection; cell degeneration disease; neurodegeneration; ds;
 XX Alzheimer's disease; schizophrenia; immune disease; inflammatory disease.
 XX Homo sapiens.
 XX FR2819824-A1.
 XX 26-JUL-2002.
 XX 23-JAN-2001; 2001FR-00000899.
 XX 23-JAN-2001; 2001FR-00000899.
 XX (MOLE-) MOLECULAR ENGINES LAB SA.
 XX Telerman A, Amson R, Tuijnder M, Susini L;
 XX WPI; 2002-610803/66.
 XX New nucleic acid implicated e.g. in tumor suppression, useful for
 XX diagnosis of tumors, viral infection and cellular degeneration and for
 XX drug screening.
 XX Claim 1; Page 141; 623pp; French.
 XX The present invention relates to novel human nucleic acid sequences (I).
 XX

CC The present sequence is one such nucleic acid sequence. Expression of (I)
CC are implicated in tumour suppression or reversion and apoptosis and viral
CC resistance. (I) are useful as probes or primers for detecting,
CC identifying, measuring and/or amplifying nucleic acid sequences, as
CC antisense reagents and for recombinant production of polypeptides. (I),
CC polypeptides (II) encoded by (I), vector containing (I), cells containing
CC these vectors and antibodies (Ab) against (II) are all useful for
CC treatment/prevention of viral, tumour and cell degeneration diseases
CC (especially neurodegeneration, such as Alzheimer's disease and
CC schizophrenia). Analysing the expression of (I) is also useful for
CC diagnosis and/or prognosis of such diseases. Transgenic animals carrying
CC (I) are used for studying the aetiology of these diseases (also immune
CC and inflammatory diseases). Note: In the present specification, SEQ ID 1
CC to 2280 are claimed in Claim 1, however only SEQ ID 1 to 2270 are shown
CC in the specification
XX
SQ Sequence 421 BP; 125 A; 77 C; 118 G; 96 T; 0 U; 5 Other;

Query Match 14.28; Score 285.8; DB 6; Length 421;
Best Local Similarity 87.68; Pred. No. 2.2e-66;
Matches 311; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 173 CTTTACACGAGGCGGAACTACTTGGACCAAAAGCGAGGACATGGAATGGCAACCA 232
DB 59 CTTTACACGAGGCGGAGACCACTTGGACCAAAAGCGAGGACATGGAATGGGAGCG 118
QY 233 CAGTGGATGAAATGACCAAAAGCAAGTATTTCATTTTGTGTTTAAAAAATGTT 292
DB 119 CAGTGGATGACATGACCAAAAGCAAGTATTTCATTTTGTGTTTAAAAAATGTT 178
QY 293 TATTGGAAGTCTTAACACAAAGATATATTCCTGATGATGTAATGTTTGTGCAAC 352
DB 179 TGTGTTGAAGTGCTCAGCACAAAGCAATAGCTCCTAGTAATGTTACTTGGTGCAGC 238
QY 353 ATGAATGGGAAAGACCAAGGCTGGCAGTGCATGTAATTTGGAGGAAGGACTTTA 412
DB 239 ATGAATGGGAAAGACCAAGGCTGGCAGTGCATGTAATTTGGAGGAAGGACTTTA 298
QY 413 GTCAAGCTCAAGGAAATGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGG 472
DB 299 GTCAACCTCAAGGAAATGGTGGAGAGGCAAGTAAATGTTTACTGGAGTAGATGGTTGG 358
QY 473 TACAGCCTGTATGTGCACTAACACCACTGAAAGAAATTAACCTAAGAGAAAT 527
DB 359 TGACTGCTGTATGTTCATCACTAACACCACTGAAAGAAATTAACCTAAGAGAAAT 413

RESULT 15

ABQ94724
ID ABQ94724 standard; DNA; 423 BP.

AC ABQ94724;

XX 28-OCT-2002 (first entry)

DE Tumour suppression-related oligonucleotide #375.

XX Tumour; cytostatic; antiviral; neuroprotective; neurotropic; neuroleptic;
KW tumour suppression; tumour reversion; apoptosis; viral resistance; human;
KW viral infection; cell degeneration disease; neurodegeneration; ds;
KW Alzheimer's disease; schizophrenia; immune disease; inflammatory disease.

XX Homo sapiens.

XX FR2819824-A1.

XX 26-JUL-2002.

XX 23-JAN-2001; 2001FR-00000899.

XX 23-JAN-2001; 2001FR-00000899.

XX (MOLE-) MOLECULAR ENGINES LAB SA.

XX

PI Telerman A, Amson R, Tuijnder M, Susini L;

XX WPI; 2002-610803/66.

XX New nucleic acid implicated e.g. in tumor suppression, useful for
PT diagnosis of tumors, viral infection and cellular degeneration and for
PT drug screening.

XX Claim 1; Page 125-126; 623pp; French.

XX The present invention relates to novel human nucleic acid sequences (I).
CC The present sequence is one such nucleic acid sequence. Expression of (I)
CC are implicated in tumour suppression or reversion and apoptosis and viral
CC resistance. (I) are useful as probes or primers for detecting,
CC identifying, measuring and/or amplifying nucleic acid sequences, as
CC antisense reagents and for recombinant production of polypeptides. (I),
CC polypeptides (II) encoded by (I), vector containing (I), cells containing
CC these vectors and antibodies (Ab) against (II) are all useful for
CC treatment/prevention of viral, tumour and cell degeneration diseases
CC (especially neurodegeneration, such as Alzheimer's disease and
CC schizophrenia). Analysing the expression of (I) is also useful for
CC diagnosis and/or prognosis of such diseases. Transgenic animals carrying
CC (I) are used for studying the aetiology of these diseases (also immune
CC and inflammatory diseases). Note: In the present specification, SEQ ID 1
CC to 2280 are claimed in Claim 1, however only SEQ ID 1 to 2270 are shown
CC in the specification
XX

SQ Sequence 423 BP; 125 A; 74 C; 124 G; 99 T; 0 U; 1 Other;

Query Match 14.18; Score 284.8; DB 6; Length 423;
Best Local Similarity 89.28; Pred. No. 4.2e-66;
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGCGGAACTACTTGGACCAAAAGCGAGGACATGGAATGGCAACCA 232
DB 71 CTTTACACGAGGCGGAGACCACTTGGACCAAAAGCGAGGACATGGAATGGGAGCG 130

QY 233 CAGTGGATGAAATGACCAAAAGCAAGTATTTCATTTTGTGTTTAAAAAATGTT 292
DB 131 CAGTGGATGACATGACCAAAAGCAAGTATTTCATTTTGTGTTTAAAAAATGTT 190

QY 293 TATTGGAAGTCTTAACACAAAGATATATTCCTGATGATGTAATGTTTGTGCAAC 352
DB 191 TGTGTTGAAGTGCTCAGCACAAAGCAATAGCTCCTAGTAATGTTTGTGTCGAGC 250

QY 353 ATGAATGGGAAAGACCAAGGCTGGCAGTGCATGTAATTTGGAGGAAGGACTTTA 412
DB 251 ATGAATGGGAAAGACCAAGGCTGGCAGTGCATGTAATTTGGAGGAAGGACTTTA 310

QY 413 GTCAAGCTCAAGGAAATGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGG 472
DB 311 GTCAACCTCAAGGAAATGGTGGAGAGGCAAGTAAATGTTGTAAGTAGATGGTTGG 370

QY 473 TACAGCCTGTATGTGCACTAACACCACTGAAAGAAATTAACCTAAGAGAAAT 516

DB 371 TGACTGCTGTATGTTCACTAACACCACTGAAAGAAATTAACCTAAGAGAAAT 414

Search completed: January 22, 2005, 11:49:54

Job time : 956.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 19, 2005, 16:54:30 ; Search time 164.5 Seconds
(without alignments)
8723.907 Million cell updates/sec

Title: US-10-069-056-12

Perfect score: 2019

Sequence: 1 agcgtggaatgcttactc.....agccgtgaactggactaa 2019

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*

5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1040.4	51.5	5049	1	US-08-336-345-1
2	1040.4	51.5	5049	1	US-08-336-345-2
3	1040.4	51.5	5049	2	US-08-647-655-1
4	1040.4	51.5	5049	2	US-08-647-655-2
5	107.8	5.3	4680	1	US-08-254-358-1
6	107.8	5.3	4680	1	US-08-475-391-1
7	107.8	5.3	4680	2	US-08-709-609-1
8	107.8	5.3	4680	5	PCT-US95-07178-1
9	107.8	5.3	4681	4	US-09-807-802A-18
10	107.8	5.3	4910	2	US-08-331-384-2
11	107.8	5.3	4910	2	US-08-836-087-2
12	107.8	5.3	4910	3	US-09-246-320-2
13	107.8	5.3	4910	3	US-09-546-738-2
14	107.8	5.3	7214	4	US-09-438-268-1
15	107.8	5.3	7557	4	US-09-770-315-3
16	107.8	5.3	8151	4	US-09-438-268-2
17	107.8	5.3	8179	4	US-09-438-268-5
18	107.8	5.3	8698	4	US-09-770-315-2
19	103	5.1	9339	4	US-09-532-594B-12
20	103	5.1	1197	4	US-09-532-594B-13
21	103	5.1	1611	4	US-09-532-594B-14
22	103	5.1	1872	4	US-09-532-594B-3
23	103	5.1	1872	4	US-09-532-594B-15
24	103	5.1	4767	4	US-09-532-594B-1
25	98.2	4.9	969	4	US-09-807-802A-10
26	98.2	4.9	1200	4	US-09-807-802A-8
27	98.2	4.9	1641	4	US-09-807-802A-6

28	98.2	4.9	1872	4	US-09-807-802A-4	Sequence 4, Appli
29	98.2	4.9	4683	4	US-09-807-802A-19	Sequence 19, Appl
30	98.2	4.9	4718	4	US-09-807-802A-1	Sequence 1, Appli
31	64.8	3.2	7218	1	US-08-232-463-14	Sequence 14, Appl
32	60	3.0	1141	4	US-09-806-708B-22	Sequence 22, Appl
33	45.2	2.2	1141	4	US-09-806-708B-22	Sequence 1, Appli
34	40.8	2.0	1664976	4	US-08-916-421B-1	Sequence 1, Appli
35	40.8	2.0	1664976	4	US-09-692-570-1	Sequence 3, Appli
36	40.2	2.0	277	3	US-09-007-005-3	Sequence 3, Appli
37	40.2	2.0	277	3	US-09-244-796-3	Sequence 3, Appli
38	39.4	2.0	580073	4	US-08-545-528D-1	Sequence 1, Appli
39	38.6	1.9	116592	4	US-09-818-512-3	Sequence 3, Appli
40	38.4	1.9	832	4	US-09-621-976-2813	Sequence 2813, Ap
41	38.4	1.9	1367	4	US-09-270-767-14265	Sequence 14265, A
42	38.4	1.9	6328	3	US-08-913-832A-1	Sequence 1, Appli
43	38.4	1.9	6328	4	US-09-249-181A-1	Sequence 1, Appli
44	38.4	1.9	6328	4	US-09-158-707-1	Sequence 1, Appli
45	38.4	1.9	6475	4	US-09-620-312D-325	Sequence 325, App

ALIGNMENTS

RESULT 1
US-08-336-345-1
; Sequence 1, Application US/08336345
; Patent No. 5814510
; GENERAL INFORMATION:
; APPLICANT: Parrish, Colin R.
; APPLICANT: Gruenberg, Allen
; APPLICANT: Carmichael, Leland E.
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,345
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Gordon, Jennifer
; REGISTRATION NUMBER: 30753
; REFERENCE/DOCKET NUMBER: 7937-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Parvovirus
; US-08-336-345-1

Query Match 51.5%; Score 1040.4; DB 1; Length 5049;
Best Local Similarity 71.4%; Pred. No. 1.9e-292;
Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;

Oy 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGAGCAACCAACTGTTAAAGGAAAA 60

Db 273 ATGCTGGCAACCAAGTATCTAGGAGAGTTATGAGGAGTAAATGTTTAAAGAAACAT 332
Qy 61 AGTAACCAAGAGTGTCTCATTTGTTTAAATAAGAAATGTTCAACTGAATGGAAAA 120
Db 333 GCAGAAATGAAGCAATTTTGGTTGTTTAAATGTGACAACTGTCACAACTAAATGGAAAG 392
Qy 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCAGAGGACGAGCTGAAATCTTTACAA 180
Db 393 GATGTTGCTGGAACCACTATACCAACCAATTCAAATGAAGCTTAACATCTTTAATT 452
Qy 181 CGAGGACCGAAACTACTTGGGACCAAGC-----GAGGACATGGAATGGGAAACACA 234
Db 453 AGAGGAGCAAAAAGCAATGATCAAAACCGAAGAGAAATGGACTGGGAATCGGAA 512
Qy 235 GTCGATGAATGACCAAAAAAGCAAGTATTCATTTTGGATCTTTGGTTTAAAAAATGTTTA 294
Db 513 GTTGATAGTCTCGCAAAAAAGCAAGTACAACTTTTGTATGCAATTAATTAATAAATGTCCT 572
Qy 295 TTTGAAGTGTCTTAACAACAAGAAATATATTTCTGGTGATGTTAATTTGTTTGTGCAACAT 354
Db 573 TTTGAAGTCTTTGTTCTAAAAATATAGAACCAATGAATGTGTTTGGTTTATTCACAT 632
Qy 355 GAATGGGAAAGACCAAGCTGCACTGCCATGTACTAATTTGGAGGAAAGACTTTAGT 414
Db 633 GAATGGGAAAGATCAAGGCTGSCATTTGTCTTTTACTTTCATAGTAAGAACTTACAA 692
Qy 415 CAAGCTCAAGGGAAATGCTGGAGAGGCAACTAAATGTTTACTTGGACAGATGTTGGTA 474
Db 693 CAAGCAACTGGTAAATGGCTACCGACAGCAAAATGAATGATTTGGAGTAGATGGTGGTG 752
Qy 475 ACAGCTGTAAATGTGCAACTAACACCAAGTGAAGAAATTAACCTAAGAGAAATAGCAGAA 534
Db 753 ACTCTTTGTTGCGTAACTTTAACACCAACTGAAGATTAAGCTCAGAGAAATTCAGAA 812
Qy 535 GACAATGAGTGGTGTACTTACTTACTTATTAAGCATTAAGCAAAACCAAAAAAGACTATACC 594
Db 813 GATAGTGAATGGGTGACTATATTAACATACAGACATTAAGCAAAACCAAAAAAGACTATGTT 872
Qy 595 AAGTGTGCTCTTTTGGAAACATGATGCTTACTTATTTTAACTTAAAAAGAAATTAAGC 654
Db 873 AAATGGTTCATTTTGGAAATATGATAGCATATTAATTTTAAACAAAGAAATTAATGTC 932
Qy 655 ACTAGTCCACCAAGAGACGGAGGTATTTTCTTAGCAGTGACTCTGCTGCGAAACCTAAC 714
Db 933 CACAT---GACAAAAGAAAGTGGCTATTTTAAAGTACTGATCTGCTTGGAAATTTAAC 989
Qy 715 TTTTAAAGAGCGGAGCGCATCTAGTGAAGCAACTATACACTGATGACATGCGGCCA 774
Db 990 TTTATGAAGTATCAAGA CAGACAAATTTGTGAGCACTTTTACACTGAACAAATGAACCA 1049
Qy 775 GAAACGGTTGAAACCAAGTAAACCACTGCGAGGAAACTAAGCGCGCAGAAATTCAAACT 834
Db 1050 GAAACGGTTGAAACCAAGTGA CAGACAGGAAACCAAGCGCGGAGAAATTCAAACT 1109
Qy 835 AAAAAAGAAAGTCTTATTAATAAAGTACATTTAAGAGCTGGTGCATAAAGAGTAACCTCA 894
Db 1110 AAAAAAGAAAGTCAATCAATGATTTTTCGGGACTTGGTTAGTAAAGAGTAAACATCA 1169
Qy 895 CCAGAGACTGGATGATGATGACGACGACAGATTTACATTTGAATGATGGCTCAACAGGT 954
Db 1170 CCTGAAGACTGGATGATGTTTACACCAAGATGTTATATTGAATGATGGCAGCAACACGGA 1229
Qy 955 GGAGAAACCTGCTGAAAAATACCTAGAGATTTGTACACTTAACCTTAGCGCAGAACCAAA 1014
Db 1230 GTTGAATCTTTTAAAAAATACACTTGAATTTGTACTTTGACTTTAGCAAGACAAAA 1289
Qy 1015 ACAGCAATTTGACTTAAATTTTAGAAAAAGCTGAAACCAAGCAAACTTAACCTTTTCACTG 1074
Db 1290 ACAGCAATTTGAAATTAATTTAGAAAAAGCAGATAATACTAAACTTAACCTTTGATCTT 1349
Qy 1075 CTTGACACAGAACTCGAATTTTGTCTTTTTCATGCTGGCACTATGTTAAAGTTTGC 1134

Db 1350 GCAAAATCTAGAACATGTCAAATTTTTAGAAATGCACGGATGGAATTCGATTAAGTTTGT 1409
Qy 1135 CATGCTATTTGCTGTGTTTAAACAGACAGAGGAGGCAAAAGAAATGCTGTTTATTTTCAT 1194
Db 1410 CAOGCTATAGCATGTGTTTAAATAGACAGAGTGGTAAAGAAATACAGTTCITTTTTTCAT 1469
Qy 1195 GACCCAGCAGCAGCAGCAAAATCTATTTATTTGCA CAAGCCATAGCA CAAGCAGTTGGCAAT 1254
Db 1470 GGACCAAGAGTACAGGAAATCTATCATTTGCTCAAGCCATAGCACAAAGCTGTGGTAAAT 1529
Qy 1255 GTTGGTGTCTATTAATCAGCCCAATGTAACTTTCCATTTTAAATGACTGTACCAACAAGAAC 1314
Db 1530 GTTGGTGTCTATTAATCAGCAAAATGTAAATTTTCCATTTTAAATGACTGTACCAATAAAT 1589
Qy 1315 TTCAATTTGGGTAGCAAGAGCTGCTAACTTTTGACAGCAAGTAAACCAAGTTTAAAGCCATT 1374
Db 1590 TTAATTTGGATGAAGAGCTGCTAACTTTTGGTCAACAAGTAACTCAATTTAAAGCAATC 1649
Qy 1375 TGCTTGGTCAAACTATTCGCATTTGATCAAAAAGAAAGGAGGAGCAACAGATTTGAACCA 1434
Db 1650 TGTCTGGACAAACAAATTAAGATTTGATCAAAAAGGTAAGGAAAGTAAGCAAAATTTGAACCA 1709
Qy 1435 ACACCACTCATCATGACCAAAATGAGAACATTTACAGTGTGTCAGAAATAGCTCGAAGAA 1494
Db 1710 ACTCCAGTAAATATGACAACTAATGAATTAACAAATTTGAGAAATTTGATGTAAGAA 1769
Qy 1495 AGACCAAGAACACACTCAACCAATCAGAGACAGAAATGCTTAACTTCACTTAACACATACC 1554
Db 1770 AGACCTGAACATACACAAACCAATAAGAGACAGAAATGTTGAACATTTAAGTTAGTATAG 1829
Qy 1555 TTGCTTGGTGAATTTGGTGTGACAAAATTAAGAAATTTGGCCCATGATTTGCTTGGTTG 1614
Db 1830 CTTCCAGGAGACTTTGGTGTGATTAAGAAAGAAATGGCCCTTTAATATGTCATGTTA 1889
Qy 1615 GTAAAGATGTTTACCATCTTACATGGCAAGCTACTGTCTTAAATCGGCGCAAGTTTCTT 1674
Db 1890 GTTAAACATGTTTGAATCAACATGGCTTACTATACATCATTTGGGGAAGAAAGTACCA 1949
Qy 1675 GATTGCTCAGAAAACCTGGGCGAGCCAAAGGTGCCAACTCTTATAAATTTACTAGTTTCG 1734
Db 1950 GAATGGATGAAACCTGGGCGAGCCTTAAATACAAAGAGGTATAAATTCACAGGTTGC 2009
Qy 1735 GCAGCTCACCATTACGACACCGAAAGTACGCTTCTAGCCAGACATATGCACTAACT 1794
Db 2010 AAAGACT---TAGAGACACAAGCGGCAAGCAATCTCTAGAGTCAAGACCAAGTTCTAACT 2066
Qy 1795 CCATTTGATCGATCTCGAGGACCTGGCTTTTAGAGCTTTGGAGCACACCAAAATACCTCT 1854
Db 2067 CCTGACTCCGAGCTAGTGGACCTTGGACTGGAACTGGAGTACTCCAGATACGCT 2126
Qy 1855 GTTGGCGGCACTGCAGAAACCCAGAACACTTGGGGAAGCTGGTTCCAAAGCTGCAAGAT 1914
Db 2127 ATTGAGAAACTGCAAAATCAACAAATCAACCAACTTGGCGTGTACTCA-----CAAGAC 2180
Qy 1915 GGTCAACTGAGCCCACTTGGTCAGAGATCGAGGAGGATTTGAGAGGCTGCTCGGTGCG 1974
Db 2181 GTGCAAGCGAGTCCGAGTGTGCTCCGAAATAGAGCAGACCTGAGAGCCATCTTTACTCT 2240
Qy 1975 GAACCGTTGAAGAAAGACTTTCAAGCA 2000
Db 2241 GAACAATTTGAAGAAAGATTTTCAGCA 2266

RESULT 2

US-08-336-345-2

; Sequence 2, Application US/08336345

; Patent No. 5814510

; GENERAL INFORMATION:

; APPLICANT: Parrish, Colin R.

; APPLICANT: Gruenberg, Allen

; APPLICANT: Carmichael, Leland E.

; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine

; NUMBER OF SEQUENCES: 2

Qy	1615	GTAAAGAATGGTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGCAAAAGTTCCT	1674
Db	1890	GTTAAACATGGTTTTGAATCAACCATGGCTAACTATACACATCATTTGGGAAAGTACCA	1949
Qy	1675	GATTGGTCAGAAAACCTGGGGGGAGCCAAAGGTGCCAACTCCTATAAATTTACTAGGTTGG	1734
Db	1950	GAATGGGATGAAAACCTGGGGAGCCTTAAATATACAGAAGGTATAAATTCACCAGGTTGC	2009
Qy	1735	GCAGCTCACCATTCAGACACCGCAAAAGTACGCTCTCAGCCAGBACTATGCACCTAACT	1794
Db	2010	AAAGACT--TAGAGACACAAGCGGCAAGCAATCCTCAGAGTCAGACCAGTTCTAACT	2066
Qy	1795	CCACTTCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAAAATACCTCT	1854
Db	2067	CCTCTGACTCGGACGTAAGTGAGACCTTGCACCTGGAACCTGGAGTACTCCAGATACGCTT	2126
Qy	1855	GTTGCGGGCACTGCAGAAAACCCAGAACACTGGGGAGCTGGTTCAAAGCCTGCCAAGAT	1914
Db	2127	ATTGCGAANAATGCAATCAACATCAACCAACTTGGCGTTACTCA-----CAAGAC	2180
Qy	1915	GGTCAACTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTCTCGGTGGG	1974
Db	2181	GTGCAAGCGAGTCCGACGTGGTCCGAAATAGAGGCAGACCTGAGAGCCATCTTTACTTCT	2240
Qy	1975	GAACCGTTGAAGAAAGACTTCAGCGA	2000
Db	2241	GAACAATTTGAAGAAGATTTTCAGA	2266

[illegible]

Qy	1015	ACAGCATTTGACTTAATTTT	TAGAAAAAGCTGAAC	CCAGCAAACTAA	CCCAACTTTT	CACTG	1074		
Db	1290	ACAGCATTTGAATTAAT	TACTTGA	AAAAAGCAGATAAT	TACTAACTAACT	TTTGATCTTT	1349		
Qy	1075	CCTGACACAAGAACCT	CGAGAAATTTTTT	TGCTTTT	CATGGCTGGAACTAT	GTAAAGATTGCG	1134		
Db	1350	GCAAAATTC	TAGAACATGTC	AAATTTTTT	TAGAATGCAC	GGATGGATTTGAATTTGT	1409		
Qy	1135	CATGCTAATTTGCTGT	TTTTTAA	CAGACAGACAGGAGG	CAAAAGAAATGCTG	TTTTATTTCAT	1194		
Db	1410	CACGCTATAGCATGT	TTTTTAA	ATAGACAAGGTGGT	TAAGAAATACAGAT	TCTTTTTTCAT	1469		
Qy	1195	GGACGACGACACAGG	CAAAATCTATT	TATTTG	CAACAAGCCATAG	CACAAGCAGTTGGCAAT	1254		
Db	1470	GGACCAAGTACAGG	AAATCTATCT	TGCTC	AGCCATAGCA	CAAGCTGTGGGTAAAT	1529		
Qy	1255	GTTGGTTGCTATAAT	GCAGCAATGT	AAAACTTTT	CCATTTAAT	GACTGTACCAACAAGAAC	1314		
Db	1530	GTTGGTTGTTAT	ATATGACG	AAATGT	AAAAATTTTCCATTTAAT	GACTGTACCAATAAAAAAT	1589		
Qy	1315	TTGATTTGGGTAG	AAAGAGCTGGT	TAACTTT	TGGACAGCAAGT	TAAACCAAGTTTTAAAGCCATT	1374		
Db	1590	TTAATTTGGATT	GAAGAGCTGGT	TAACTTT	TGGTC	CAACAGTTAATCAATTTAAAGCAATC	1649		
Qy	1375	TGCTCTGGTCAAACT	ATATCG	CAATTTGAT	TCAAAAAGG	AAAAAGGCAAGCAAA	CAGATTGAACCA	1434	
Db	1650	TGTTCTGGAC	AAAACAAAT	TAGAAATTTGAT	TCAAAAAGGTTAA	AGGAAGTAAAGCAAA	TTGAACCA	1709	
Qy	1435	ACACGACTCAT	GTACCA	CAAAATGAGAA	CAATTTAC	TAGTGGTCAGAA	TAGGCTGGGAAGAA	1494	
Db	1710	ACTCCAGTAAT	TATGACA	CACTAAT	TGA	AAAAATATAACAATTTG	TGAAGATTTGGATGTGAAGAA	1769	
Qy	1495	AGACCAAA	CAACACTCA	ACCAATCAG	AGACAGAAATGCTT	TAACTTCACTTAA	CAACATACC	1554	
Db	1770	AGACCTGAA	CA	TACACAA	CCAAATAG	AGACAGAAATGTTGA	CACTAATAGTTAGTTAGTAAAG	1829	
Qy	1555	TTGCTCGTGAC	TTTGTGTTT	GGTTGACA	AAAAATGAAT	GGCCCAATGATTTGTGCTGGTTG	1614		
Db	1830	CTTCCAGGAGACT	TTTGTGTTT	GGTTGAT	TAAAGAA	GAATTTGGCTTTTAAATAT	TGTGCATGGTTA	1889	
Qy	1615	GTAAGAAATGTT	ACCAATCT	TACCATCG	CAAGCTTACTG	CTAAATGGGGCAAGTTCCCT	1674		
Db	1890	GTTTAA	CAATGGTTT	TGAATCA	ACATGGCTTAACT	TATACATCAT	TATGGGAAATGATACCA	1949	
Qy	1675	GATTTGT	CAGAAAACT	GGGGAGCC	AAAAAGTGC	CAACTCCTAT	TAAATTTACTAGGTTGCG	1734	
Db	1950	GAATGGGAT	GAAAACT	GGGGAGCC	TAAAAATACA	GAAGGTAT	TAAATTTCAACAGGTTGC	2009	
Qy	1735	GCACGCT	CAACATTC	ACGACAC	CGAAAGTAC	GCCTCTCAG	CCGACAGACTATGCATTA	1794	
Db	2010	AAAGACT	---	TAGAGA	CACAAGCGG	CAAGCAATCCT	CAGAGTCAAGACCAAGTTCTTA	2066	
Qy	1795	CCACTTC	GATCGGATCT	CGAGGACCT	GGCTTTTAG	AGCCTTTAG	AGCCTTTGGAGCA	CACAAAATACTCCCT	1854
Db	2067	CCTCTGACT	CCGGACG	TAGTGGAC	CTTTGCACT	TGGAA	CCGTGGAGTACTCC	CAGATACGCT	2126
Qy	1855	GTTCCGGGCACT	GCAGAAA	CCAGAAC	CACTGGGGA	AGCTGGTTCC	AAAGCCTGCCAAGAT	1914	
Db	2127	ATTGCAGAAA	CTGCAAA	TCAACAA	TCAAACTT	CGCGTTACTCA	-----CAAAGAC	2180	
Qy	1915	GGTCAA	CTGAGCCCCAA	CTTGGT	CAGAGATCG	AGGAGGATTT	TGAGAGCGGTGCTTC	GGTGCG	1974
Db	2181	GTGC	AGCGAGTCCGA	CGTGGT	CCGAAT	TAGGCG	CAACCTG	AGAGCCATCTTTACTTCT	2240
Qy	1975	GAACCGTT	GAGAAAG	ACTTC	ACGCA	2000			
Db	2241	GAACAA	TTTGAAG	AGATTTT	TCGAGA	2266			

Patent No. 5885585
GENERAL INFORMATION:
APPLICANT: Parrish, Colin R.
APPLICANT: Gruenberg, Allen
APPLICANT: Carmichael, Leland E.
TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/647,655
FILING DATE: On Even Date Herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gordon, Jennifer
REGISTRATION NUMBER: 30,753
REFERENCE/DOCKET NUMBER: 7937-008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5049 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Parvovirus
US-08-647-655-2

Query Match	51.5%;	Score 1040.4;	DB 2;	Length 5049;
Best Local Similarity	71.4%;	Pred. No. 1.9e-252;		
Matches 1432;	Conservative	0;	Mismatches 556;	Indels 18; Gaps 4;
Qy	1	ATGCGTGGAAATGCTTACTCTGTGAAGTTTTGGGAGCAACCACTGGTTAAAGGAAAA	60	
Db	273	ATGCTGGCAACCMGNTATCTGAGGAAGTTATGGAGGGAGTAAATTTGGTTTAAAGAAACAT	332	
Qy	61	AGTAACCAAGGAAGTGTCTCATTTGTTTTTAAAAAATGAAAATGTTCAACTGAATGGAAAA	120	
Db	333	GCAGAAAAATGAAGCATTTTCGTTGTTTTTAAATGTGACACGCTCCAACTAAATGGGAAG	392	
Qy	121	GATATCGGATGGATAGTTACAAAAAGAGCTGCAGGAGGACGAGCTGAATCTTTTACAA	180	
Db	393	GATGTTGCGTCGGAACAACTATACCAAAACCAATTCAAAATGAAGAGCTTAACATCTTTAATT	452	
Qy	181	CGAGGAGCGGAAACTACTCTGGGACCAAGC-----GAGGACATGGAAATGGGAAACCA	234	
Db	453	AGAGGACACAAACAGCAATGGATCAAAACCGAAGAAGAAATGGACTGGGAATCGAA	512	
Qy	235	GTGGATGAATGACCAAAAAGCAAGTATTCATTTTTGATTTCTTTGGTTTAAAAAATGTTTA	294	
Db	513	GTTTGATAGTCTGCGCAAAAAGCAAGTACAAACTTTTGATGCATTAATTTAAAAAATGTC	572	
Qy	295	TTTGAAGTGTCTTAACACAAAAGATATATTTTCCTGGTGATGTAAATCGGTTGTGCAACAT	354	
Db	573	TTTGAAGTCTTTGTTCTTAAAAATATAGAACCAAAATGAATGTGTTTGTATTCACAT	632	
Qy	355	GAATGGGAAAAGAACCAAGCTGGCACTGCCATGTACTAAATGGAGGAAAGGACTTTAGT	414	
Db	633	GAAATGGGAAAAGATCAAGCTGGCATGTTCATGTTTTACTTTCATAGTAGAAGCTTACAA	692	

QY 415 CAAGCTCAAGGGAATGGTGGAGAGCAACTAAATGTTTACTTGGAGCAGATGGTTGTA 474
Db |||||
QY 693 CAAGCAACTGGTAAATGGCTAGCCAGACAAATGAATATGTTATGGGAGATGGTTGGTG 752
Db |||||
QY 475 ACAGCCTGTAATGTGCAACTAACACACAGCTGAAAGAAATTAACCTAGAGAAATAGCAGAA 534
Db |||||
QY 753 ACTCTTTGTCGGTAAACTTTAACACCACTGAAAGAAATTAAGCTCAGAGAAATTCAGAA 812
QY 535 GACAAATGAGTGGTGTACTCTACTTACTTATAAGCATAGCAACCAACCAAAAGACTATACC 594
Db |||||
QY 813 GATAGTGAATGGTGTACTATATTAACATACAGACATAAGCAACCAAAAGAAATTAAGTGT 872
QY 595 AAGTGTGTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAGAAATTAAGC 654
Db |||||
QY 873 AAAATGGTTCATTTTGGAAATATGATAGCATATTTTAACTAAAGAAATTAAGTGT 932
QY 655 ACTAGTCCACCAAGAGAGCGAGCTATTTCTTAGCAGTACTCTGCTGGAACCTAAC 714
Db |||||
QY 933 CACAT---GACAAAGAAAGTGGCTATTTTAACTAAAGTACTGATTTCTGTTGGAAATTTAAC 989
QY 715 TTTTAAAGAGCGAGCGCCATCTAGTGAGCAAACTATATACACTGATGACATGGCGCCA 774
Db |||||
QY 990 TTTATGAAGTATCAGACAGACAAATTTGTGAGCAGCTTTTACACTGAACAAATGAAACCA 1049
QY 775 GAAACGGTTGAAACCAAGTAAACCTGCGCAGGAACTAAGCGCGCAGAAATTTCAAAC 834
Db |||||
QY 1050 GAAACCGTTGAAACCAAGTAAACCTGCGCAGGAACTAAGCGCGGAGAAATTTCAAAC 1109
QY 835 AAAAAAGAAAGTTCATTTAAACCTTACACTTAAAGAGCTGGTGCAATAAAGAGTAACTCA 894
Db |||||
QY 1110 AAAAGGAAGTGTCAATCAATGATGTTTTCGGGACTTGGTTAGTAAAGAGTAACTCA 1169
QY 895 CCAGAGACTGGATGATGAGCGAGCAGCAGTACTTGAATGATGGCTCAACAGGT 954
Db |||||
QY 1170 CTTGAAGACTGGATGATGTTTACACACAGATAGTTATTTGAAATGATGGCAGCAACAGGA 1229
QY 955 GGAGAAACCTGCTGAAAAATACCTAGATGATTTGTACACTAACTCTAGCCAGAACCA 1014
Db |||||
QY 1230 GGTGAAATCTTTTAAATAATCACTTGAATTTGTACTTTGACTTTAGCAGAACAA 1289
QY 1015 ACAGCAATTGACTTAAATTTTAGAAAAAGCTGAAACCAAGCAAACTAACCACTTTCACTG 1074
Db |||||
QY 1290 ACAGCAATTGAAATTAATTAACCTGAAAGAGCAGATAATACTAACTAACTTTGATCTT 1349
QY 1075 CTTGACACAGAACTGCAAAATTTTCTTTTCACTGCTGGAATGTTTAAAGTTTG 1134
Db |||||
QY 1350 GCAAAATCTAGAACATGCAAAATTTTGAATGACCGATGGAAATGGATTAAGTTTGT 1409
QY 1135 CATGCTATTTGCTGTTTAAACAGACAAAGGAGGCAAAAGAAATGCTGTTTATTTTCAT 1194
Db |||||
QY 1410 CACGCTATAGCAATGTTTAAATAGACAGGTGGTAAAGAAATACAGTTCTTTTTCAT 1469
QY 1195 GGACAGCCAGCAGCAGCAATCTTATTTAGCAAGCCATAGCAGCAAGCTTTGGCAAT 1254
Db |||||
QY 1470 GGACAGCAAGTACAGGAAATCTATCATTTGCTCAAGCCATAGCAGCAAGCTTTGGTAA 1529
QY 1255 GTTGGTGTCTATATGAGCAAAATGAACTTTTCACTTTTAACTGCTGCTGCAACAGAAC 1314
Db |||||
QY 1530 GTTGGTGTCTATATGAGCAAAATGAAATTTTCACTTTTAACTGCTGCTGCAACAGAAC 1589
QY 1315 TTGATTTGGGTAGAGAGCTGGTAACTTTTGGACAGCAAGTAAACCAAGTTTAAAGCCATT 1374
Db |||||
QY 1590 TTAATTTGGATGAGAGCTGGTAACTTTTGGTCAACAAAGTTAATCAATTTAAAGCAATC 1649
QY 1375 TGCTGTGCTCAAACTATTTGCAATGATCAAAAGAAAGCAAGCAACAGCTTTGACCA 1434
Db |||||
QY 1650 TGTTCTGACAAACAAATTTAGAAATTTGATCAAAAGGTTAAAGGAAATTTGAAACCA 1709
QY 1435 ACACAGTCTATGACCAACCAAAATGAGAACTTACAGTGGTCAAGATAGGCTGCGAAGAA 1494
Db |||||
QY 1710 ACTCCAGTATATGACAACTAATGAAATATACAAATTTGTGAGAAATTTGATGTGAAGAA 1769
Db |||||

QY 1495 AGACCAGAACACACTCAACCAATCAGACAGAGAAATGCTTAAACATTCATCTTAACATACC 1554
Db |||||
QY 1770 AGACTGAAACATACACAAACCAATTAAGACAGAGAAATGTTGAACATTTAGTTAGTATGTAAG 1829
Db |||||
QY 1555 TTGCTGTGACTTTTGGTTGGTTGACAAAATGAATGCGCCATGATTTGCTTGGTTG 1614
Db |||||
QY 1830 CTTCCAGGAGACTTTTGGTTGGTTGATAAAGAAATGGCCCTTTAATATGTCATGGTTA 1889
QY 1615 GTAAAGAAATGGTTTCAATCTACCAATGCGCAAGCTACTGTGCTTAAATGGGGCAAAATTCCT 1674
Db |||||
QY 1890 GTTAAACATGGTTTGAATCAACATGGCTACTATACATCATTTGGGAAAAGTACCA 1949
QY 1675 GATGGTTCAGAAAACCTGGCGGAGCAAGAGTGGCAACTCTCTATATAATTTACTAGTTGCG 1734
Db |||||
QY 1950 GAATGGATGAAAACCTGGCGGAGCTAAATAACAAAGAGGTATAAATTCACAGGTTGC 2009
QY 1735 GCAGCTCACCATTACGACACCGAAAGTACGCTCTCAGCCAGAACTATGCTACTTAAT 1794
Db |||||
QY 2010 AAAGACT---TAGAGACACAAAGCGCAAGCAATCTCAGAGTCAAGACCAAGTTCTAAT 2066
QY 1795 CCAGTTGCATCGATCTCGAGGACCTGGCTTTTAGAGCTTTGGAGCAGCAACCAATACTCT 1854
Db |||||
QY 2067 CTTCTGACTCCGAGCGTAGTGGACCTTGCACTGGAAACCGTGGAGTACTCCAGATAGCCT 2126
QY 1855 GTTGGCGGCACTGCAAGAAACCCAGAACACTGGGGAGCTGGTTCCAAAGCCTGCCAAGAT 1914
Db |||||
QY 2127 ATTGCAGAACTGCAAAATCAACAATCAACCAACTTGGCGTTACTCA-----CAAAGAC 2180
QY 1915 GGTCACTGAGCCCACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGGTCTTCGCTGGG 1974
Db |||||
QY 2181 GTGCAAGCAGTCCGAGCTGGTCCGAAATAGAGGAGCAGCTGAGAGCCATCTTACTTCT 2240
QY 1975 GAACCGTTTGAAGAAAGACTTCAGCGA 2000
Db |||||
QY 2241 GAACAATTGGAAGAGATTTTCGAGA 2266
Db |||||

RESULT 5
US-08-254-358-1
; Sequence 1, Application US/08254358
; Patent No. 5658785
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/254,358
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5658785and, Greta B.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31975
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4680 base pairs
;


```
;      nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-254-358-1

Query Match          5.3%; Score 107.8; DB.1; Length 4680;
Best Local Similarity 53.4%; Pred. No. 9e-21;
Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

QY      1168   GGCAAGAAGAAATGCTGTTTATTTCATGGACACCAGCCAGCAGCGCAAACTCTATTATTGCA 1227
        |||||
Db       1293   GGCAGAGGAACACCACATCTGGCTGTTTGGCGCTGCAACTACGGGAAGACAACATCGCG 1352

QY      1228   CAAGCCATAGCAAAAGCAGTGTGGCAATGTTGGTTGCTTATAATCGAGCCAATGTAAACTTTT 1287
        |||||
Db       1353   GAGGCCATAGCCACACACTGTGCCCTTCTACGGGTGCGTAACCTGGACCAATGAGAACTTT 1412

QY      1288   CCATTTAATAGCTGTACCACACAGAAGCTTGATTTGGTAGAAGAGCTGTTAACTTTGGA 1347
        |||||
Db       1413   CCCTTCAACCACTGTGTGCAAGAAGTGGTGATCTGGTGGAGGAGGGGAAGATGACCGCC 1472

QY      1348   CAGCAAGTAACACAGTTTAAAGGCCATTGCTCTGGTCAAACATAATTCGCAATTGATCAAAA 1407
        |||||
Db       1473   AAGTCTGAGTCTGGCCAAAGCCATTCTCGGAGGAAGCAAGGTGCGCTGGACCAAGAA 1532

QY      1408   GSAAGAGGCGAGCAAAACAGATTGAAACCAACACCAAGTCATCATGACCAAAATGAGAACATT 1467
        |||||
Db       1533   TGCAAGTCTCTGGGCCACAGATAGACCCGACTCCCGTGATCGTCACCTCCAACACCAACATG 1592

QY      1468   ACAGTGGTCAAGATAGCTCGGAGAAGAACACAGACACACTCAACCAATCAGAGACAGA 1527
        |||||
Db       1593   TGGCGCGTGATTGACCGGAACTCAACGACCTTCAACACACAGCAGCGCGTTGCAAGACCCGG 1652

QY      1528   ATGCTTAACATTTCATCTAACACATACCTTGCCTGGTGACTTTGGTTGGTTGACAAAAAT 1587
        |||||
Db       1653   AUGTTCAATTTGAACTCACCGCCGCTCTGGATCATGTACTTTGGGAAGGTTCACCAAGCAG 1712

1588   GAA 1590
    |||
1713   GAA 1715
```

RESULT 6
US-08-475-391-1
; Sequence 1, Application US/08475391
; Patent No. 5786211
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and
; METHOD OF INVENTION: Methods
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,391
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/254,358
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5786211and, Greta E.

```

; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31975
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-475-391-1

Query Match          5.3%; Score 107.8; DB 1; Length 4680;
Best Local Similarity 53.4%; Pred. No. 9e-21;
Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

QY 1168 GGC AAAAGAAATGCTGTTTATTTTCATGTGACACGACGACGACGCAAAATCTATTATTGCA 1227
DB      |||
QY 1293 GGC AAGAGGAACACCATCTGGCTGTTTGGGCGCTGCAACTACCGGAAGACCAACATCGCG 1352
DB      |||
QY 1228 CAAGCCATAGCACAAAGCAGTTGCGCAATGTTGGTTGCTATTAATGCAGCCAATGTAAACTTT 1287
DB      |||
QY 1353 GAGGCCATAGCCACACACTGTGCCCTTCTACGGGTGCGTAAACTGGACCAATGAGAACTTT 1412
DB      |||
QY 1288 CCATTTTAATGACTGTACCAACAAGAACTTGATTTGGGTAGAAGAAGCTGGTAACTTTTGA 1347
DB      |||
QY 1413 CCCTTCAACGACTGTGTGCAAGATGGTGATCTGGTGGAGAGAGGGGAAGATGACCGCC 1472
DB      |||
QY 1348 CAGCAAGTAAACAGATTTAAAGCCATTTGCTCTGGTCAAACTATTCGCAATTGATCAAAAA 1407
DB      |||
QY 1473 AAGGTCGTGGAGTCGCCCAAAGCCATTCTCGGAGGAAGCAAGGTGCGCGTGGACCGAAA 1532
DB      |||
QY 1408 GGAAGAGGCAGCAAAAGAGATTGAACCAACACCAAGTATCATGACCACAAATGAGAACTT 1467
DB      |||
QY 1533 TGC AAGTCCTCGGCCCCAGATAGACCGGACTCCCGTGATCGTCACTTCCACACCAACATG 1592
DB      |||
QY 1468 ACAGTGGTCAGATAGGCTGCGAAGAAAGACCAAGAACACACTCAACCAATCAGAGACAGA 1527
DB      |||
QY 1593 TGGCGCGTGATTCAGCGGGAACCTCAAGCACTTCGAACACCAAGCGGTTGCAAGACCGG 1652
DB      |||
QY 1528 ATGCTTTAAACATTCATCTAAACATACCTTGGCTGGTGACTTTGGTTGGTTGACAAAAAT 1587
DB      |||
QY 1653 ATGTTCAAATTTGAACTCAACCCCGCGTCTGGATCATGACTTTGGGAAGGTCACCAACGAG 1712
DB      |||
QY 1588 GAA 1590
DB      |||
QY 1713 GAA 1715

RESULT 7
US-08-709-609-1
; Sequence 1, Application US/08709609
; Patent No. 5858775
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```

```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/709,609
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5858775and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31975
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
;
US-08-709-609-1
;
; Query Match 5.3%; Score 107.8; DB 2; Length 4680;
; Best Local Similarity 53.4%; Pred. No. 9e-21;
; Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0;
;
QY 1168 GGCARAGAAATGCTGTTTATTTTCATGGACCCAGCAGCAGCAAAATCTATTATTGCA 1227
Db 1293 GGCAGAGAGAACACCATCTGGCTGTTGGCCCTTGCACCACTACCGGAGACCAACATCGCG 1352
QY 1228 CAAGCCATAGCAGCAAGCAGTTGGCAATGTTGTTGCTATATATGACGCCAAATGTAACCTTT 1287
Db 1353 GAGCCATAGCCACACTGTGCCCTTCTACGGGTGCGTAACTGGACCAATGAGAACTTT 1412
QY 1288 CCATTTAATGACTGTACCAACAGAACTTGAATTTGGTAGAAGAGCTGTTACTTTGGA 1347
Db 1413 CCCTTCAACGACTGTGCGCAAGATGGTGATCTGTTGGAGGAGGGGAGAGATGACCGCC 1472
QY 1348 CAGCAAGTAACACAGTTTAAAGCCATTTGCTGCTGCTCAAACTATTTCGCATTGATCAAAA 1407
Db 1473 AAGTCGTGGAGTCGCGCAAGCCATTTCTGGAGGAGCAAGTGCGCGTGGACGAGAA 1532
QY 1408 GMAAAGGCGAGCAACAGATTGAACCAACACAGCTATCATGACCACAAATGAGAACATT 1467
Db 1533 TGCAAGTCTCTCGGCCAGATAGACCCGACTCCCGTGATCGTCACTTCAACACCAACATG 1592
QY 1468 ACATGTCAGAAATAGCTGCGAAGAGACCAACACACACTCAACCAATCAGAGACAGA 1527
Db 1593 TGGCCCGTGATTGACGGGAATCAACGACCTTCGAAACACAGAGCCGTTGCAAGACCG 1652
QY 1528 ATGCTTAACATTCTATCAACATACATCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1587
Db 1653 ATGTTCAAAATTGAACTACCCCGCGTCTGGATCATGACTTTGGGAAGGTCAACCAAGCAG 1712
QY 1588 GAA 1590
Db 1713 GAA 1715

```

```

RESULT 8
PCT-US95-07178-1
; Sequence 1, Application PC/TUS9507178
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606

```

```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07178
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Noland, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31975
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
;
PCT-US95-07178-1

```

```

;
; Query Match 5.3%; Score 107.8; DB 5; Length 4680;
; Best Local Similarity 53.4%; Pred. No. 9e-21;
; Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0;
;
QY 1168 GGCARAGAAATGCTGTTTATTTTCATGGACCCAGCAGCAGCAAAATCTATTATTGCA 1227
Db 1293 GGCAGAGAGAACACCATCTGGCTGTTGGCCCTTGCACCACTACCGGAGACCAACATCGCG 1352
QY 1228 CAAGCCATAGCAGCAAGCAGTTGGCAATGTTGTTGCTATATATGACGCCAAATGTAACCTTT 1287
Db 1353 GAGCCATAGCCACACTGTGCCCTTCTACGGGTGCGTAACTGGACCAATGAGAACTTT 1412
QY 1288 CCATTTAATGACTGTACCAACAGAACTTGAATTTGGTAGAAGAGCTGTTACTTTGGA 1347
Db 1413 CCCTTCAACGACTGTGCGCAAGATGGTGATCTGTTGGAGGAGGGGAGAGATGACCGCC 1472
QY 1348 CAGCAAGTAACACAGTTTAAAGCCATTTGCTGCTGCTCAAACTATTTCGCATTGATCAAAA 1407
Db 1473 AAGTCGTGGAGTCGCGCAAGCCATTTCTGGAGGAGCAAGTGCGCGTGGACGAGAA 1532
QY 1408 GMAAAGGCGAGCAACAGATTGAACCAACACAGCTATCATGACCACAAATGAGAACATT 1467
Db 1533 TGCAAGTCTCTCGGCCAGATAGACCCGACTCCCGTGATCGTCACTTCAACACCAACATG 1592
QY 1468 ACATGTCAGAAATAGCTGCGAAGAGACCAACACACACTCAACCAATCAGAGACAGA 1527
Db 1593 TGGCCCGTGATTGACGGGAATCAACGACCTTCGAAACACAGAGCCGTTGCAAGACCG 1652
QY 1528 ATGCTTAACATTCTATCAACATACATCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1587
Db 1653 ATGTTCAAAATTGAACTACCCCGCGTCTGGATCATGACTTTGGGAAGGTCAACCAAGCAG 1712
QY 1588 GAA 1590
Db 1713 GAA 1715

```

```

RESULT 9
US-09-807-802A-18
; Sequence 18, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNVFN.031USA

```

; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 4681
; TYPE: DNA
; ORGANISM: AAV-2
US-09-807-802A-18

Query Match 5.3%; Score 107.8; DB 4; Length 4681;
Best Local Similarity 53.4%; Pred. No. 9e-21;
Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0;
QY 1168 GGCAGAGAAATGCTGTTTATTTCATGGACCGCAGCAGCAATCTATTATGCA 1227
DB 1293 GGCAGAGAGAAATGCTGTTTATTTCATGGACCGCAGCAGCAATCTATTATGCA 1352
QY 1228 CAAGCCATAGCAGCAAGCAGTTGGCAATGTTGGTGTCTATATGCAGCCAAATGTAACATTT 1287
DB 1353 GAGGCCATAGCCACACTGTCCTCTACGGGTGCGTAACTGGACCAATGAGAACTTT 1412
QY 1288 CCATTAAATGACTGTACCAACAAGAACTTGATTTGGGTAGAAAGCTGTTAACTTTGGA 1347
DB 1413 CCCTTCAACGACTGTGTCGACAGATGGTGATCTGGTGGGAGGAGGAGATGACCGCC 1472
QY 1348 CAGCAAGTAAACCAAGTTTAAAGCAATTTGCTCTGGTCAAACTATTGCGCATTTGATCAAAAA 1407
DB 1473 AAGTCTGTGAGTGGCGCAAGCAATTTCTCGGAGGAAGCAAGTGGCGGTGGACCAAGAA 1532
QY 1408 GGAAGGAGGAGCAACAGATTTGACCAACCAAGTATCATGACCAATGAGAACTTT 1467
DB 1533 TGCAAGTCTCTCGGCCAGATAGACCGGACTCCCGTGTATGTCACCTCCAAACCAACATG 1592
QY 1468 ACAGTGTCTAGATAGCTCGGAGGAAGAACAGACACACTCAACCAATCAGAGACAGA 1527
DB 1593 TGGCGCGTGTATGACGGGAATCAAGCACTTCAACACAGCAGCGGTTGCAAGACCGG 1652
QY 1528 ATGCTTAAACATTCATTAACACATACCTTCGCTGGTGAATTTGGTTGGTTGCAAAAAAT 1587
DB 1653 ATGTTCAAATTTGAATCACTCCCGCGTCTGGATCATGACTTTGGGAAGGTCAACCAAGCAG 1712
QY 1588 GAA 1590
DB 1713 GAA 1715

RESULT 10
US-08-331-384-2/c
; Sequence 2, Application US/08331384
; Patent No. 5856152
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Kelley, William M.
; APPLICANT: Fisher, Krishna J.
; TITLE OF INVENTION: Hybrid Adenovirus-AAV Vector and Methods
; TITLE OF INVENTION: of Use Thereof
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P.O. Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/331,384
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: UPNG1149USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4910 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
US-08-331-384-2

Query Match 5.3%; Score 107.8; DB 2; Length 4910;
Best Local Similarity 53.4%; Pred. No. 9.3e-21;
Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0;
QY 1168 GGCAGAGAAATGCTGTTTATTTCATGGACCGCAGCAGCAATCTATTATGCA 1227
DB 1544 GGCAGAGAGAAATGCTGTTTATTTCATGGACCGCAGCAGCAATCTATTATGCA 1485
QY 1228 CAAGCCATAGCAGCAAGCAGTTGGCAATGTTGGTGTCTATATGCAGCCAAATGTAACATTT 1287
DB 1484 GAGGCCATAGCCACACTGTCCTCTACGGGTGCGTAACTGGACCAATGAGAACTTT 1425
QY 1288 CCATTAAATGACTGTACCAACAAGAACTTGATTTGGGTAGAAAGCTGTTAACTTTGGA 1347
DB 1424 CCCTTCAACGACTGTGTCGACAGATGGTGATCTGGTGGGAGGAGGAGATGACCGCC 1365
QY 1348 CAGCAAGTAAACCAAGTTTAAAGCAATTTGCTCTGGTCAAACTATTGCGCATTTGATCAAAAA 1407
DB 1364 AAGTCTGTGAGTGGCGCAAGCAATTTCTGGAGGAAGCAAGTGGCGGTGGACCAAGAA 1305
QY 1408 GGAAGGAGGAGCAACAGATTTGAAACCAACCAAGTATCATGACCAACCAATGAGAACTTT 1467
DB 1304 TGCAAGTCTCTCGGCCAGATAGACCGGACTCCCGTGTATGTCACCTCCAAACCAACATG 1245
QY 1468 ACAGTGTCTAGATAGCTCGGAGGAAGAACAGACACACTCAACCAATCAGAGACAGA 1527
DB 1244 TGGCGCGTGTATGACGGGAATCAAGCACTTCAACACAGCAGCGGTTGCAAGACCGG 1185
QY 1528 ATGCTTAAACATTCATTCATACACATACCTTCGCTGGTGAATTTGGTTGGTTGCAAAAAAT 1587
DB 1184 ATGTTCAAATTTGAATCACTCCCGCGTCTGGATCATGACTTTGGGAAGGTCAACCAAGCAG 1125
QY 1588 GAA 1590
DB 1124 GAA 1122

RESULT 11
US-08-836-087-2/c
; Sequence 2, Application US/08836087
; Patent No. 5871982
; GENERAL INFORMATION:
; APPLICANT: Trustees of University of Pennsylvania
; APPLICANT: Wilson, James M.
; APPLICANT: Kelley, William M.
; APPLICANT: Fisher, Krishna J.
; TITLE OF INVENTION: Hybrid Adenovirus-AAV Vector and
; TITLE OF INVENTION: Methods of Use Thereof
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, PO Box 457

CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836,087
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA: 08/331,384
APPLICATION NUMBER: 08/331,384
FILING DATE: 28-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNVFN.007PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4910 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cdna
US-08-836-087-2

Query Match	5.3%	Score 107.8	DB 2	Length 4910
Best Local Similarity	53.4%	Pred. No. 9.3e-21		
Matches 226	Conservative 0	Mismatches 197	Indels 0	Gaps 0
Qy	1168	GGCAAAAGAAATGCTGTTTATTTTCATGAGCAGCAGCAAAATCTATTATTGCA	1227	
Db	1544	GGCAAGAGGAACACCATCTGCTGTTTGGGCTGCAACTACCGGGAAGACCAACATCGG	1485	
Qy	1228	CAAGCCATAGCAACAAGCAGTTGGCAATGTTGGTTGCTATAATGCAGCCCATGTAACTTT	1287	
Db	1484	GAGGCCATAGCCCACTGTGCCCCTTCTACGGGTGCGTAACTGGACCAATGAGAATCTT	1425	
Qy	1288	CCATTTAATGACTGTACCAACAAGAACTTGATTTGGGTAGAAAGCTGTGTAATTTTGG	1347	
Db	1424	CCCTTCAACGACTGTGTCGACAAGATGGTGTATCTGGTGGGAGGAGGGAAGATGACGCC	1365	
Qy	1348	CAGCAAGTAACCAAGTTTAAAGCCATTTGCTCTGGTCAAACTATTTCGATTTGATCAAAA	1407	
Db	1364	AAGGTCGTGGAGTCGGCCAAAGCCATTTCTCGGAGGAAGCAAGTGCCTGGAGCAGAAA	1305	
Qy	1408	GGAAAGGCAGCAAAAGAGTTGAACCAACACACGCTCATATGACCAAAATGAGAACTT	1467	
Db	1304	TGCAAGTCCTCGGCCAGATAGACCCGACTCCCGTGATGTCACCTCCACACCAACATG	1245	
Qy	1468	ACAGTGTTCAGAAATAGGTCGGAAGAAGACAGCAACACACTCAACCAATCAGAGACAGA	1527	
Db	1244	TGCGCCGTGATTGACGGGAACTCAACGACCTTCGAAACACCAGCAGCCGTTGCAAGACCGG	1185	
Qy	1528	ATGCTTTAACATTTCACTTAACACATACCTTGCTCGTGACTTTGTTTGGTTGACAAAAT	1587	
Db	1184	ATGTTCAAAATTTGAATCTACCCGCCGCTTGGATCATGACTTTGGGAAGGTCACCAAGCAG	1125	
Qy	1588	GAA	1590	
Db	1124	GAA	1122	

RESULT 12

RESULT 12
US-09-246-320-2/c

US-09-246-320-2/C
; Sequence 2, Application US/09246320

; sequence 2, Applicant
; Patent No. 6251677

```

: GENERAL INFORMATION:
: APPLICANT: Trustees of University of Pennsylvania
: APPLICANT: Wilson, James M.
: APPLICANT: Kelley, William J.
: APPLICANT: Fisher, Krishna J.
: TITLE OF INVENTION: Hybrid Adenovirus-AAV Vector and
: TITLE OF INVENTION: Methods of Use Thereof
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Howson and Howson
: STREET: Spring House Corporate Cntr, PO Box 457
: CITY: Spring House
: STATE: Pennsylvania
: COUNTRY: USA
: ZIP: 19477
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/246,320
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/836,087
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Bak, Mary E.
: REGISTRATION NUMBER: 31,215
: REFERENCE/DOCKET NUMBER: GNVFN.007PCT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 215-540-9200
: TELEFAX: 215-540-5818
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4910 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: unknown
: MOLECULE TYPE: CDNA
: US-09-246-320-2

```

Query Match	5.3%	Score 107.8	DB 3	Length 4910
Best Local Similarity	53.4%	Pred. No. 9.3e-21		
Matches 226	Conservative 0	Mismatches 197	Indels 0	Gaps 0
Qy	1168	GGCAAGAAGAAATGCTGTTTATTTTCATGGACACGACAGCAGGCAAAATCTATTATTGCA	1227	
Db	1544	GGCAAGAGGAAACCAATCTGGCTGTTTGGGCGCTGCAACTACCGGGAAGACCAACATCGCG	1485	
Qy	1228	CAAGCCATAGCAACAAGCAGTGTGGCAATGTTGGTGTGCTATAATGCAGCCAAATGTAAACTTTT	1287	
Db	1484	GAGGCCATAGCCACACTGTGCCCTTCTACGGGTGGTAAATCTGGACCAATGAGAACTTT	1425	
Qy	1288	CCATTTAATGACTGTACCAACAAGAACTTGATTGGGTAGAAGAAGCTGTTAACTTTGGGA	1347	
Db	1424	CCCTTCACGACTGTGTGCACAAGATGGTGATCTGGTGGGAGGAGGGAAGATGACCGCC	1365	
Qy	1348	CAGCAAGTAAACGAGTTTAAAGCCATTTGCTCTGGTCAAACTATTGCGATTGATCAAAA	1407	
Db	1364	AAGTCTGTGGTGGCGCAAAAGCCATTCTCGGAGGAAGCAAGGTGGCGGTGGACCCAGAA	1305	
Qy	1408	GGAAGGCGAGCAAAACAGATTGAAACCAACACCACTCATATGACCAAAATGAGAACATT	1467	
Db	1304	TGCAAGTCTTCGGCCAGATAGACCCGACTCCCGTGATGCTGCTACCTTCAACACCAATG	1245	
Qy	1468	ACAGTGTGCAGATAGGCTCGAAGAAGAACACAGAAACACACTCAACCAATCAGAGACAGA	1527	
Db	1244	TGGCGCTGATTGACGGGAATCTCAACGACTTCGAACACCAAGCAGCGGTGCAAGACCGG	1185	
Qy	1528	ATGCTTAAACATTTCATAACACATACTTGCGCTGGTGACTTTGGTTGGTTGACAAAAT	1587	

Db 1184 ATGTTCAAAATTGAACCTACCCCGCTCTGGATCATGACTTTGGGAAGGTCACCAAGCAG 1125

Qy 1588 GAA 1590

Db 1124 GAA 1122

RESULT 13

US-09-546-738-2/c

; Sequence 2, Application US/09546738

; Patent No. 6387368

; GENERAL INFORMATION:

; APPLICANT: Trustees of University of Pennsylvania

; Wilson, James M.

; Kelley, William M.

; Fisher, Krishna J.

; TITLE OF INVENTION: Hybrid Adenovirus-AAV Vector and

; Methods of Use Thereof

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Howson and Howson

; STREET: Spring House Corporate Cntr, PO Box 457

; CITY: Spring House

; STATE: Pennsylvania

; COUNTRY: USA

; ZIP: 19477

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/546,738

; FILING DATE: 11-Apr-2000

; CLASSIFICATION: <unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/246,320

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Bak, Mary E.

; REGISTRATION NUMBER: 31,215

; REFERENCE/DOCKET NUMBER: GNVN.007PCT

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 215-540-9200

; TELEFAX: 215-540-5818

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4910 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: unknown

; MOLECULE TYPE: cDNA

; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-546-738-2

Query Match 5.3%; Score 107.8; DB 3; Length 4910;

Best Local Similarity 53.4%; Pred. No. 9.3e-21;

Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

Qy 1168 GGCAAAAGAAATGCTGTTTATTTCATGGACCGCAGCAGCAGCAATCTATTATTGCA 1227

Db 1544 GGCAAGAGGAACCACTCTGGCTGTTTGGGCTGCAACTACCGGGAAGAGCAACATCGCG 1485

Qy 1228 CAAGCCATAGCAGTGGCAATGTTGGTGTCTATATGCGCAATGTAACTTT 1287

Db 1484 GAGGCCATGCCACACTGTGCCCTTCTACGGGTGCGTAACTGGACCAATGAGAACTTT 1425

Qy 1288 CCATTAAATGACTGTACCAACAAGAACTGATTGGGTAGAGAAGCTGGTAACTTTGGA 1347

Db 1424 CCCTTCAACGACTGTGCGACAGATGGTGATCTGGTGGGAGGAGGAGATGACCGCC 1365

Qy 1348 CAGCAAGTAAACCAAGTTTAAAGCCATTTGCTCTGGTCAAACTATTGCGCATTTGATCAAAA 1407

Db 1364 AAGTCTGTGAGTCGGCCAAAGCCATTCTCGGAGGAAGCAAGGTGCGCGTGGACCAGAA 1305

Qy 1408 GGAAGAGGCGAGCAAAACAGATTGAACCAACACAGTCTATCATGACCAAAATGAGAAACATT 1467

Db 1304 TGCAAGTCTCTCGGCCAGATAGACCGGACTCCCGTGATCGTCACTCCAAACCAACATG 1245

Qy 1468 ACAGTGTGAGATAGCTCGGAAGAAAGACCAAGACACACTCAACCAATCAGAGACAGA 1527

Db 1244 TCGCCGTGATTGACGGGAACCTCAACGACCTTCGAACACACAGCAGCGCTTGAAGACCGG 1185

Qy 1528 ATGCTTAACATTCTATCTAACACATACCTTCCTCGGTGACTTTGTTGTTGTTGACAAAAAT 1587

Db 1184 ATGTTCAATTTGAATCACTCCCGCGTCTGGATCATGACTTTGGGAAGGTCACCAAGCAG 1125

Qy 1588 GAA 1590

Db 1124 GAA 1122

RESULT 14

US-09-438-268-1/c

; Sequence 1, Application US/09438268

; Patent No. 6491907

; GENERAL INFORMATION:

; APPLICANT: Rabinowitz, Joseph E.

; APPLICANT: Samulski, Richard J

; APPLICANT: Xiao, Weidong

; TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING

; FILE REFERENCE: 5470-186

; CURRENT APPLICATION NUMBER: US/09/438,268

; CURRENT FILING DATE: 1999-11-10

; EARLIER APPLICATION NUMBER: 60/107,840

; EARLIER FILING DATE: 1998-11-10

; EARLIER APPLICATION NUMBER: 60/123,651

; EARLIER FILING DATE: 1999-03-10

; NUMBER OF SEQ ID NOS: 59

; SOFTWARE: Patent in Ver. 2.0

; SEQ ID NO 1

; LENGTH: 7214

; TYPE: DNA

; ORGANISM: Virus

US-09-438-268-1

Query Match 5.3%; Score 107.8; DB 4; Length 7214;

Best Local Similarity 53.4%; Pred. No. 1.2e-20;

Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

Qy 1168 GGCAAAAGAAATGCTGTTTATTTCATGGACCGCAGCAGCAGCAATCTATTATTGCA 1227

Db 3899 GGCAAGAGGAACCACTCTGGCTGTTTGGGCTGCAACTACCGGGAAGACCAACATCGCG 3840

Qy 1228 CAAGCCATAGCAGCAGTGGCAATGTTGGTGTCTATATGCGCAATGTAACTTT 1287

Db 3839 GAGGCCATAGCCACACTGTGCCCTTCTACGGGTGCGTAACTGGACCAATGAGAACTTT 3780

Qy 1288 CCATTAAATGACTGTACCAACAAGAACTTGAATTTGGGTAGAGAAGCTGGTAACTTTGGA 1347

Db 3779 CCCTTCAACGACTGTGCGACAGATGGTGATCTGGTGGGAGGAGGAGATGACCGCC 3720

Qy 1348 CAGCAAGTAAACCAAGTTTAAAGCCATTTGCTCTGGTCAAACTATTGCGCATTTGATCAAAA 1407

Db 3719 AAGTCTGTGAGTCGGCCAAAGCCATTCTCGGAGGAAGCAAGGTGCGGTGGACCAGAA 3660

Qy 1408 GGAAGAGGCGAGCAAAACAGATTGAACCAACACAGTCTATCATGACCAAAATGAGAAACATT 1467

Db 3659 TGCAAGTCTCTCGGCCAGATAGACCGGACTCCCGTGATCGTCACTCCAAACCAACATG 3600

Qy 1468 ACAGTGTGAGATAGCTCGGAAGAAAGACCAAGACACACTCAACCAATCAGAGACAGA 1527

Db 3599 TGGCCGTGATTGACGGGAACCTCAACGACCTTCGAACACACAGCAGCGCTTGAAGACCGG 3540

Qy 1528 ATGCTTAACATTCTATCAACATACCTTGCCTTGGTGTGACTTTGTTGTTGTTGACAAAAAT 1587

Db 3539 ATGTTCAAAATTGAACATCCCGCGTCTGGATCATGACTTTGGGAAGGTACCAAGCAG 3480
Qy 1588 GAA 1590
Db 3479 GAA 3477

RESULT 15
US-09-770-315-3
; Sequence 3, Application US/09770315
; Patent No. 6429001
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; TITLE OF INVENTION: Recombinant AAV Packaging Systems
; FILE REFERENCE: 20263-501
; CURRENT APPLICATION NUMBER: US/09/770,315
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,536
; PRIOR FILING DATE: 2000-01-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 7557
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: recombinant DNA
US-09-770-315-3

Query Match 5.3%; Score 107.8; DB 4; Length 7557;
Best Local Similarity 53.4%; Pred. No. 1.2e-20;
Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0;
Qy 1168 GGCARAAGAATGCTGTTTATTTCATGGACCCAGCCAGCAGCAAAATCTATTATTGCA 1227
Db 1321 GGCAAGAGGAACACCATCTGGCTGTTTGGCCCTGCAACTACCGGGAAGACCAATCGCG 1380
Qy 1228 CAAGCCATAGCACAAGCAGTTGGCAATGTTGGTTGCTTATATGAGCAGCCAAATGTAAACTTT 1287
Db 1381 GAGGCCATAGCCCACTGTGCCCTTCTACGGGTGCGTAACTGGACCAATGAGAACTTT 1440
Qy 1288 CCATTTAATGACTGTACCAAGAACTTGATTTGGGTAGAAGAGCTGTTAACTTTGGA 1347
Db 1441 CCTTCAACGACTGTGTCAGAGATGATGATCTGTTGGAGGAGGGGAAAGATGACCGCC 1500
Qy 1348 CAGCAAGTAAACCAAGTTTAAAGCCATTTGCTCTGTTCAAACTATTGCAATTGATCAAAA 1407
Db 1501 AAGTCTGGAGTCCGCCAAGCCATTCTCGGGAAGCAGAGTGGCGGTGGACCAAAA 1560
Qy 1408 GGAAGGCGCAACAAAGATTGAAACCAACCAAGTCAATGATGACCAAAATGAGAACTT 1467
Db 1561 TGCAAGTCTCGGCCAGATAGACCCGACTCCCGTATGCTGCTCACTCCAAACCAACATG 1620
Qy 1468 ACAGTGTGAGATAGCTGGGAAGAAAGACCAAGACCAACTCAACCAATCAGAGACAGA 1527
Db 1621 TGCCTGCTGATGTCAGCGGAATCAACGACCTTCGAACACCAAGCCGCTTGCAGACCGG 1680
Qy 1528 ATGCTTAACATTCATCAACACATACCTTGCTGGTGCATTTGGTTGGTTGACAAAAAT 1587
Db 1681 ATGTTCAAAATTGAATCAACCCCGCGTCTGGATCATGACTTTGGGAAGGTACCAAGCAG 1740
Qy 1588 GAA 1590
Db 1741 GAA 1743

Search completed: January 19, 2005, 17:05:56
Job time : 171.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2005, 11:50:09 ; Search time 1072 Seconds
(without alignments)

10821.780 Million cell updates/sec

Title: US-10-069-056-12

Perfect score: 2019

Sequence: 1 agcggcggaagtcttactc.....agccgctgaactggactaa 2019

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1771	87.7	5121	US-10-647-111A-1	Sequence 1, Appli
2	941.4	46.6	5075	US-10-361-002-3	Sequence 3, Appli
3	941.4	46.6	5075	US-10-361-004-3	Sequence 3, Appli
4	307.4	15.2	374	US-10-466-894-1376	Sequence 1276, Ap
5	285.8	14.2	421	US-10-466-894-430	Sequence 430, App
6	284.8	14.1	423	US-10-466-894-375	Sequence 375, App
7	284.8	14.1	464	US-10-466-894-370	Sequence 370, App
8	284.8	14.1	473	US-10-466-894-367	Sequence 367, App
9	284.8	14.1	486	US-10-466-894-358	Sequence 358, App
10	284.8	14.1	491	US-10-466-894-428	Sequence 428, App
11	284.2	14.1	420	US-10-466-894-438	Sequence 438, App
12	283.8	14.1	343	US-10-466-894-388	Sequence 388, App

13	283.8	14.1	343	18	US-10-466-894-390	Sequence 390, App
14	283.8	14.1	403	18	US-10-466-894-433	Sequence 433, App
15	283.2	14.0	497	18	US-10-466-894-369	Sequence 369, App
16	280.4	13.9	340	18	US-10-466-894-383	Sequence 383, App
17	278.2	13.8	342	18	US-10-466-894-416	Sequence 416, App
18	273.8	13.6	420	18	US-10-466-894-432	Sequence 432, App
19	269.6	13.4	324	18	US-10-466-894-391	Sequence 391, App
20	263.4	13.0	339	18	US-10-466-894-1275	Sequence 1275, Ap
21	261.6	13.0	424	18	US-10-466-894-429	Sequence 429, App
22	163.8	8.1	468	18	US-10-466-894-524	Sequence 524, App
23	162.2	8.0	451	18	US-10-466-894-519	Sequence 519, App
24	162.2	8.0	457	18	US-10-466-894-518	Sequence 518, App
25	162.2	8.0	465	18	US-10-466-894-525	Sequence 525, App
26	162.2	8.0	469	18	US-10-466-894-520	Sequence 520, App
27	162.2	8.0	472	18	US-10-466-894-531	Sequence 531, App
28	162.2	8.0	515	18	US-10-466-894-444	Sequence 444, App
29	162.2	8.0	516	18	US-10-466-894-445	Sequence 445, App
30	162.2	8.0	516	18	US-10-466-894-451	Sequence 451, App
31	162.2	8.0	516	18	US-10-466-894-458	Sequence 458, App
32	162.2	8.0	517	18	US-10-466-894-462	Sequence 462, App
33	162.2	8.0	530	18	US-10-466-894-442	Sequence 442, App
34	162.2	8.0	530	18	US-10-466-894-450	Sequence 450, App
35	161	8.0	516	18	US-10-466-894-447	Sequence 447, App
36	160.6	8.0	530	18	US-10-466-894-456	Sequence 456, App
37	159	7.9	468	18	US-10-466-894-523	Sequence 523, App
38	133.4	6.6	247	18	US-10-466-894-471	Sequence 471, App
39	117	5.8	1386	9	US-09-792-630-32	Sequence 32, Appl
40	117	5.8	1386	10	US-09-953-351-32	Sequence 32, Appl
41	117	5.8	1386	13	US-10-080-376-32	Sequence 38, Appl
42	117	5.8	1386	14	US-10-082-671-38	Sequence 38, Appl
43	117	5.8	1386	14	US-10-097-100-32	Sequence 32, Appl
44	117	5.8	1386	15	US-10-023-208-32	Sequence 32, Appl
45	115.4	5.7	1884	9	US-09-792-630-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1

US-10-647-111A-1
; Sequence 1, Application US/10647111A
; Publication No. US20040209240A1
; GENERAL INFORMATION:
; APPLICANT: MALERBA, MADDALENA
; TITLE OF INVENTION: ANTI-NEOPLASTIC VIRUSES
; FILE REFERENCE: 604-691
; CURRENT APPLICATION NUMBER: US/10/647,111A
; CURRENT FILING DATE: 2003-08-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 1
; LENGTH: 5121
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Parvovirus H1
; OTHER INFORMATION: with promoter P4 and left hairpin from WMV (pH1)
US-10-647-111A-1

Query Match	87.7%	Score	1771	DB	18	Length	5121
Best Local Similarity	92.3%	Pred. No.	0				
Matches	1864	Conservative	0	Mismatches	155	Indels	0
Gaps	0						
QY	1	ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA	60				
DB	282	ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA	341				
QY	61	AGTAACACGAGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAA	120				
DB	342	AGTAACACGAGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGAAAA	401				
QY	121	GATATCGGATGGAATAGTTTACAAAAAAGAGCTGCAGGAGGAGCTGAAATCTTTTACAA	180				

Db 402 GATATCGGATGGATAGTTTACAAAAGAGCTGCAGGAGCAGAGCTGAAATCTTTTACAA 461
Qy 181 CGAGGAGCGGAACTACTCTGGGACCAAAGCAGGACATGGAATGGGAAACCAAGTGGAT 240
Db 462 CGAGGAGCGGAACTACTCTGGGACCAAAGCAGGACATGGAATGGGAAACCAAGTGGAT 521
Qy 241 GAAATGACCAAAAGCAAGTATTCATTTTGGTTTCTTTGTTTAAATAATGTTTATTGAA 300
Db 522 GAAATGACCAAAAGCAAGTATTCATTTTGGTTTCTTTGTTTAAATAATGTTTATTGAA 581
Qy 301 GTGCTTTAAACAAAGATATATTTCTGCTGATGTTTAAATGTTGTCACCAATGAATGG 360
Db 582 GTGCTTAAACAAAGATATATTTCTGCTGATGTTTAAATGTTGTCACCAATGAATGG 641
Qy 361 GGAAGACCAAGCTGGCAGTCGCCATGTATTAATTTGGAGGAAGACATTTAGTCAAGCT 420
Db 642 GGAAGACCAAGCTGGCAGTCGCCATGTATTAATTTGGAGGAAGACATTTAGTCAAGCT 701
Qy 421 CAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCC 480
Db 702 CAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGGTAAACAGCC 761
Qy 481 TGTAAATGTGCAACTAACACAGCTGAAAGAAATTAATTAAGAGAAATAGCAGAGCAAT 540
Db 762 TGTAAATGTGCAACTAACACAGCTGAAAGAAATTAATTAAGAGAAATAGCAGAGCAAT 821
Qy 541 GAGTGGGTTACTTACTTACTTATAGCATATAGCAACCAAAAGACATATACCAGTGT 600
Db 822 GAGTGGGTTACTTACTTACTTATAGCATATAGCAACCAAAAGACATATACCAGTGT 881
Qy 601 GTTCTTTTGGAAACATGATGCTTACTTATTTTAACTAAAAAGAAATTAAGCACTAGT 660
Db 882 GTTCTTTTGGAAACATGATGCTTACTTATTTTAACTAAAAAGAAATTAAGCACTAGT 941
Qy 661 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTACTCTGGCTGGGAAAACTAACTTTTAA 720
Db 942 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTACTCTGGCTGGGAAAACTAACTTTTAA 1001
Qy 721 AAAGAGGCGAGGCGCCATCTAGTGAGCAACATATACATGATGACATGCGGCCAGAAACG 780
Db 1002 AAAGAGGCGAGGCGCCATCTAGTGAGCAACATCTGTATATCTGATGAGATGAACCCAGAAACG 1061
Qy 781 GTTGAACACACAGTAAACCTGCGCAGGAACTAAAGCGCGCAGAAATTCAAACTAAAAAA 840
Db 1062 GTGAGACCAAGTACCACTGACAGAGAGCTTAAGCGCGCAGAAATTCAAACTAGAGAG 1121
Qy 841 GAAAGTTTCTATTAAAACTACACTTAAAGAGCTGGTGCAATAAAGAGTAACCTCACCAGAG 900
Db 1122 GAGGTCTCGATTAAACCACTCAAGAGTTGGTACATATAAAGAGTAACCTCACCAGAA 1181
Qy 901 GACTGGATGATGACGAGCAGAGTTCATTAATGATGCTCAACAGGTTGGAGAA 960
Db 1182 GACTGGATGATGACGAGCAGAGTTCATTAATGATGCTCAACAGGTTGGAGAA 1241
Qy 961 AACTGCTGAAAAATACGCTAGAGATTGTACACTACTCTAGCCAGAACCAAAACAGCA 1020
Db 1242 AACTGCTTAAAAATACCTAGAGATCTGTACACTGACTCTAGCAGAACCAAAACAGCC 1301
Qy 1021 TTTGACTTAAATTTTAAAAAGAGTGAACCCAGCAAACTAAACCACTTTTCACTGCTGAC 1080
Db 1302 TTTGACTTAAATTTTAAAAAGAGTGAACCCAGCAAACTAGCCAACTTTTCCATGGCTAGC 1361
Qy 1081 ACAAGACCTGCAGAAATTTTGGTTTCTTTTCTGCTGGGAACTATGTTAAAGTTTGCATGCT 1140
Db 1362 ACCAGAACCTGTAGAAATCTTTGCTGAGCATGGCTGGAACTATATTAAGTCTGCCATGCC 1421
Qy 1141 ATTTGCTGTGTTTAAACAGACAGAGGAGCAAAAGAAATGCTGTTTATTTTATTCATGGACCA 1200
Db 1422 ATCTGTTGTGCTGAATAGACAGAGGAGGCAAAAGAAACACTGTGCTCTTTTCCAGGACCA 1481
Qy 1201 GCCAGCACAGCAATCTATTATTGCAACAGCCATAGCACAGCAAGTGGCAATGTTGGT 1260

Db 1482 GCCAGCACAGGCAAACTATTATTGCACAAGCCATAGCAACAGCAGTGGTAAATGTTGGT 1541
Qy 1261 TGCTATAATGACGCAAACTGTAAACCTTTTCCATTTAATGACTGTACCAACAAGAACTTGATT 1320
Db 1542 TGTTACAATGCTGCCAATGTGNACTTTCATTTAATGACTGTACCAACAAGAACTTGATT 1601
Qy 1321 TGGGTAGAAAGAGCTGGTAACTTTTGGACAGCAAGTAAACCCAGTTTTAAAGCCATTTGCTCT 1380
Db 1602 TGGGTGGAAGAGCTGGTAACTTTTGGCCAGCAAGTAAACCCAAATTCAAAAGCTATTGTTCT 1661
Qy 1381 GGTCAAACTATTTCGCAATGATCAAAAGGAAGGACGCAACACAGATTGAACCAACACCA 1440
Db 1662 GGCCAAACCATACGCAATGATCAAAAGGAAAGGACGCAACACAGATTGAACCAACACCA 1721
Qy 1441 GTCATCATGACCAACAAATGAGAACTTACAGTGGTCAGAAATAGGCTGCGAAGAAAGACCA 1500
Db 1722 GTTATTATGACCAACCAACGAGNAACATTACCGTGGTTAGNATAGGCTGTGAGGAAAGACCA 1781
Qy 1501 GAACACACTCAACCAATCAGAGACAGAATGCTTAACTTCAATCATCTAACACATACCTTGCT 1560
Db 1782 GAACACACTCAACCAATCAGAGACAGAATGCTTCAACATTCACCTGACACGTAACCTACCT 1841
Qy 1561 GGTGACTTTGGTTTGGTTGACAAATAATGAATGSCCCATGATTTGTCTTGGTTGGTAAAG 1620
Db 1842 GGTGACTTTGGTTTGGTTGATTAAGCAGCAATGGCCCTCTGATCTGCTTGGTTGGTGAAG 1901
Qy 1621 AATGGTTACCAATCTACCATGGCAAGCTACTGTCTGCTAAATGGGGCAAAAGTTCTCTGATTGG 1680
Db 1902 AATGGTTACCAATCTACCATGGCTTTGTTTACTGTCTTAAATGGGGCAAAAGTTCTCTGATTGG 1961
Qy 1681 TCAGAAAACTGGGCGGAGCCAAAAGTGCCAACTCTTATATAATTTACTAGTTTGGGCACGC 1740
Db 1962 TCAGAGGACTGGGCGGAGCCGAAAGCTAGACACTCTCTATAAAATTCGCTAGGTTCAATGGCC 2021
Qy 1741 TCACCAATTCACCAACCAAGAGAGTACGCTCTCAGCCAGAACTATGACATACTTCCACTT 1800
Db 2022 TCACCAATCTCTGACTCTCGAGAAAGTACGCTCTCAGCCAAACTACGCTTCTTACTCCACTT 2081
Qy 1801 GCATCGGATCTCGAGGACCTGGCTTTTAGAGCCTTGGAGCACCACCAAACTACTCTGTTGGC 1860
Db 2082 GCATCGGACCTTGGGACCTAGCTCTAGAGCCTTGGAGCACCACCAAACTACTCTGTTGGC 2141
Qy 1861 GGCATGCGAGAAACCCAGAACTGGGAGAGCTGGTTCCAAAGCCTGCCAAGTGTCTAA 1920
Db 2142 GGCATGCGAGCAAGCCAAACCACTGGGAGGCTGGTTCCACAGCCTGCCAAGGTTCTCAA 2201
Qy 1921 CTGAGCCCAACTTGGTTCAGAGATCGAGGAGGATTTGAGAGCCTGCTTCCGTCGGGACCG 1980
Db 2202 CGAGGCCCAACTGCTGCTCCAGATCGAGGCGGATTTGAGAGCTTGTCTTCAGTCAAGAACAG 2261
Qy 1981 TTGAAGAAAGACTTTCAGCGAGCCGCTGAACCTTGGACTAA 2019
Db 2262 TTGGAGGCGACTTCAACGAGGAGCTGACCTTGGACTAA 2300

RESULT 2

US-10-361-002-3
; Sequence 3, Application US/10361002
; Publication No. US20040170954A1
; GENERAL INFORMATION:
; APPLICANT: Clearant, Inc.
; APPLICANT: McKenney, Keith
; APPLICANT: Gillmeister, Lidja
; APPLICANT: Marlowe, Kristina
; APPLICANT: Armistead, David
; TITLE OF INVENTION: Pathogen Inactivation Assay
; FILE REFERENCE: CI-0043
; CURRENT APPLICATION NUMBER: US/10/361,002
; CURRENT FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 5075

; TYPE: DNA
; ORGANISM: Porcine parvovirus
US-10-361-002-3

Query Match 46.6%; Score 941.4; DB 17; Length 5075;
Best Local Similarity 70.4%; Pred. No. 2.8e-237;
Matches 1305; Conservative 0; Mismatches 536; Indels 12;

QY	4	GCTGGAAATGCTTACTCTGATGAAGTTTTCGGAGCAACCAACTGGTTAAAGGAAAAAAGT	63
Db	298	GCGGGAACAACCTTACTCGGAAGAGGTACTAAAAGCTACCAACTGGCTTCAAGATAATGCT	357
QY	64	AACAGAGAGTGTTCTCATTTGTTTTTAAAAATGTAATGTTCAACTGAATGGAAAAAGAT	123
Db	358	CAAAAAGAGCATTTCTTATGTATTTTAAAAACA AAAAAGTCAATCTTAATTGGAAAAGAA	417
QY	124	ATCGGATGGAAATAGTTTACAAAAAGAGCTGCGAGGAGCAGCTGAAATCTTTTCAACGA	183
Db	418	ATTGCTTGGAAATAACTACAACAAAGATACAACAGATCGGAAATGATAAAACCTTACAAAGA	477
QY	184	GGAGCGGAAACTACTTGGGACCAAGCGAGGACATGGAATGGGNAACCAAGTGGATGAA	243
Db	478	GGAGCGAACAACATCATGGGACCAAGGCAACAGACATGGAATGGGAATCAGAAAATCGACAGC	537
QY	244	ATGACCAAAAAGCAAGTAATTCATTTTTGATTTCTTTGTTAAAAAATGTTTATTTGAAGTG	303
Db	538	CTCAGAAAACGGCAAGTACTGATTTTTTGACTCTCTGTTAAAAAATGTCTCTTTGAAGGT	597
QY	304	CTTAAACA CAAAGAAATATATTTTCCTGGTGATGTTTAATTTGGTTTGTGCAACATGAATGGGA	363
Db	598	ATATTGCCAAAAGAACCTTAAGTCCAAGTAGCTGCTACTTGGTTTCATACAGCATGAACATGGT	657
QY	364	AAAGACCAAGCTGGCACTGCCATGTACTTAATTTGGAGGAAAAGNCTTTAGTCAAGCTCAA	423
Db	658	CAAGATACCTGGCTATCATCTGCCATGTACTACTAGGTGGNAAGGCTTTACAAACAGCAATG	717
QY	424	GGCAATGGTGGAGAGCGCAACTAAATGTTTACTTGGAGCAGATGTTTGGTTAAACAGCCCTGT	483
Db	718	GGAAATGGTTTCAGAAAAACAATTAACAATTTATGGAGTAGTAGTTTAATATGCAATGC	777
QY	484	AATGTGCAACTAACACACAGCTGAAAAGAAATTAACCTAAGAGAGAAATPAGCAGAAAGACAATGAG	543
Db	778	AAAGTACTCTTAACACACAGTTGAAAAGAAATAAAATTTAAGGGAAATTAGCAGAGGATGGTGAG	837
QY	544	TGGGTTACTCTACTTACTTATNAGCATACGCAACCAAAAAGACTATACCAAGTGTGTT	603
Db	838	TGGGTATCGCTACTTAACCTACACTCAACAACAAACTAAAAAACAATATACAAAATGACT	897
QY	604	CTTTTTGAAAAATGATTGCTTACTATTTTTTAACTAAAAAGAAAAATAGCACTAGTCCA	663
Db	898	CATTTTGGAAATATGATTGCTTACTTCTCTTAATTAANAAGAAAGACAACCT-----	951
QY	664	CCAAGAGACGGAGGCTATTTCTTTCAGTGACTCTCGCTGGAAAACTAATCTTTTAAAAA	723
Db	952	GAAAGAGAGCATGGATATTATCTCAGCTCAGATTCCTGGCTTCATGACAAATTTCTTAAAA	1011
QY	724	GAAGCGCGGCCACTAGTTCAGCAAACTATACACTGATGACATCGCGCCAGCAAAACGGTT	783
Db	1012	GAAGCGGAGAGACAATTAGTCAGTCACCTATTTTCTTGAAGCAATAAACCTGAAACCTGTG	1071
QY	784	GAAACCAACAGTAACCACTGCGCAGAAAATAAGCGCGCAGAAATTCACAACTAAAAAGAA	843
Db	1072	GAAACAACGGTTACTACAGCTCAGAGAGCCCAAGAGGCGAGNATACAAACAANAAGAA	1131
QY	844	GTTTCTATTAAAACTACACTTTAAAGAGCTGGTGCATAAAAAGAGTAACTCTCACAGAGGAC	903
Db	1132	GTAAGCATAAAAATGCACAAATAAGAGACTTGGTTAATAAAAAAGATGTACTAGCATAGAAGAC	1191
QY	904	TGGATGATGATGCCAGCAGACAGTTTACATTTGAATGATGCTCAACAGCTGGAGAAAC	963
Db	1192	TGGATGATGATGACAGATCCAGACAGTTATATAGAAATGATGGCTCAACCGGAGGAGAAAT	1251
QY	964	CTGCTGAAAAAATACGCTTAGAGATTGTACACTAACTCTAGCCAGAACCAAAAACAGCATTT	1023

RESULT 3

RESULTS 3
US-10-361-004-3

US-10-381-004-3
: Sequence 3. Application US/10361004

; Publication No. US20040170981A1

GENERAL INFORMATION:

APPLICANT: Clearant, Inc.

APPLICANT: McKenney, Keith

APPLICANT: Gillmeister, Lidja

APPLICANT: Marlowe, Kristina

APPLICANT: Armistad, David

TITLE OF INVENTION: Real-Time Polymerase Chain Reaction Using Large Target Amplicons

FILE REFERENCE: CI-0042

; CURRENT APPLICATION NUMBER: US/10/361,004


```

; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; TITLE OF INVENTION: AND THEIR USE AS MEDICINES
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1276
; LENGTH: 374
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 291, 300, 310, 316, 347, 348, 370, 371, 374
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-1276

Query Match          15.2%; Score 307.4; DB 18; Length 374;
Best Local Similarity 88.2%; Pred. No. 1.8e-70;
Matches 329; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY      174  TTTACAACGAGGAGCGGAAACTACTTGGGACCAAGCGAGACATGGAATGGGAAACCCAC 233
         |||||
DB       1   TTTACAACGAGGCGGAGACCATTGGGACCAAGCGAGACATGGAATGGGAGAGCGC 60

QY      234  AGTGGATGAAATGATCCAAAAAGCAAGTATTTCATTTTGGTTTAAAAAATGTTT 293
         |||||
DB       61  AGTGGATGATGATGACCAAAAAAGCAAGTATTTATTTTGGTTTAAAGAAAGTGT 120

QY      294  ATTGAAGTCTTAAACACAAAGAATATATTCCTGGTGATGTTAAATGGTTTGTGCAACA 353
         |||||
DB       121 GTTTGAAGTGCTCAGCAACAAAGAACATAGCTCTAGTAATGTTACTTGGTTCTGCGCAGCA 180

QY      354  TGAATGGGGAAGACCAACAGCTGGCACTGCGCATGTACTAATTCGAGAAAGGACTTTAG 413
         |||||
DB       181 TGAATGGGGAAGACCAACAGCTGGCACTGTCATGCTGATGAGGCAAGGACTTTAG 240

QY      414  TCAAGCTCAAGGGAAATGTTGGGAAGGCAACTAAATGTTTACTGGAGCAGATGTTGGT 473
         |||||
DB       241  TCAACCTCAAGGAAATGTTGGGAAGGCACTAAATGTTGCTGAGTANATGTTTGN 300

QY      474  ACAGCCTGTAATGTGCAACTAACACACAGCTGAAAGAAATTAACCTAGAGAAATAGCAGA 533
         |||||
DB       301  GACTGCTGNAATGTNCACTTAACACAGCTGAAAGAAATTAACCTGNNAGAAATAGCAGA 360

QY      534  AGACAATGAGTGG 546
         |||||
DB       361  GGACAGTGANNGG 373

RESULT 5
US-10-466-894-430
; Sequence 430, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; TITLE OF INVENTION: AND THEIR USE AS MEDICINES
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23

```


Db 372 TGACTGCTGTAAATGTTCAACTAACACGAGTGAACAAAAA 415
|||||

RESULT 9

US-10-466-894-358
; Sequence 358, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 358
; LENGTH: 486
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 413, 428, 435, 448, 457, 462
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-358

Query Match 14.1%; Score 284.8; DB 18; Length 486;
Best Local Similarity 89.2%; Pred. No. 1.9e-64;
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGAGCGGAACTACTTTGGACCAAAAGCGAGGACATGGAATGGGAACCA 232
Db 54 CTTTACACGAGGAGCGGAGCACCTTTGGACCAAAAGCGAGGACATGGAATGGGAGCG 113

QY 233 CAGTGGATGAATGACCAAAAGCAAGTATTCATTTTGGATTTCTTTGGTTAAAAAATGTT 292
Db 114 CAGTGGATGACATGACCAAAAGCAAGTATTTATTTTGGATTTCTTTGGTTAAAGAGTGT 173

QY 293 TATTTGAAGTGCTTAACACAAAGATATTTCTCTGTGATGTTAATTTGTTTGTGCAAC 352
Db 174 TGTTGAAGTGCTCAGCACAAAGAACATAGCTCCTAGTATGTTACTTGTGTCGAGC 233

QY 353 ATGAATGGGAAAGACCAAGGCTGGCACTGCCATGTACTAAATTTGGAGGAAAGACTTTA 412
Db 234 ATGAATGGGAAAGACCAAGGCTGGCACTGTCTATGCTGTATGTTGGAGGCAAGACTTTA 293

QY 413 GTCAAGCTCAAGGAAATGTTGGAGAAAGCAACTAAATGTTTACTGGAGCAGATGTTGG 472
Db 294 GTCAACCTCAAGGAAATGTTGGAGAAAGCAAGTAAATGTTTACTGGAGTAGATGTTGG 353

QY 473 TAACAGCTGTATGTGCAACTAACACAGCTGAAGAAATTA 516
Db 354 TGACTGCTGTAAATGTTCAACTAACACGAGTGAACAAAAA 397
|||||

RESULT 10

US-10-466-894-428
; Sequence 428, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent

; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 428
; LENGTH: 491
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 12, 421, 441
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-428

Query Match 14.1%; Score 284.8; DB 18; Length 491;
Best Local Similarity 89.2%; Pred. No. 2e-64;
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGAGCGGAACTACTTTGGACCAAAAGCGAGGACATGGAATGGGAACCA 232
Db 62 CTTTACACGAGGAGCGGAGCACCTTTGGACCAAAAGCGAGGACATGGAATGGGAGCG 121

QY 233 CAGTGGATGAATGACCAAAAGCAAGTATTCATTTTGGATTTCTTTGGTTAAAAAATGTT 292
Db 122 CAGTGGATGACATGACCAAAAGCAAGTATTTATTTTGGATTTCTTTGGTTAAAGAGTGT 181

QY 293 TATTTGAAGTGCTTAACACAAAGATATTTCTCTGTGATGTTAATTTGTTTGTGCAAC 352
Db 182 TGTTGAAGTGCTCAGCACAAAGAACATAGCTCCTAGTATGTTACTTGTGTCGAGC 241

QY 353 ATGAATGGGAAAGACCAAGGCTGGCACTGCCATGTACTAAATTTGGAGGAAAGACTTTA 412
Db 242 ATGAATGGGAAAGACCAAGGCTGGCACTGTCTATGCTGTATGTTGGAGGCAAGACTTTA 301

QY 413 GTCAAGCTCAAGGAAATGTTGGAGAAAGCAAGTAAATGTTTACTGGAGCAGATGTTGG 472
Db 302 GTCAACCTCAAGGAAATGTTGGAGAAAGCAAGTAAATGTTTACTGGAGTAGATGTTGG 361

QY 473 TAACAGCTGTATGTGCAACTAACACAGCTGAAGAAATTA 516
Db 362 TGACTGCTGTAAATGTTCAACTAACACAGCTGAAGAAATTA 405
|||||

RESULT 11

US-10-466-894-438
; Sequence 438, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 438

QY	234	AGTGGATGAAATGACCAAAAGCAAGTATTCAATTTTGTATCTTTGGTTAAAAATGTTT	293
DB	61	AGTGGATGACATGACCAAAAGCAAGTATTCAATTTTGTATCTTTGGTTAAAAATGTTT	120
QY	294	ATTTGAAGTGTCTTAACAAAGAATATATTTCTCTGGTGATGTTAAATTTGGTTTGTGCAACA	353
DB	121	GTTCGAAGTGTCTAGCAAAAGCAATAGCTCTCTAGTAAATTTACTTTGGTTCTGTCAGCA	180
QY	354	TGAATGGGAAAGACCAAGCTGGCACTGCAATGTACTTAATTTGGAGGAAGACATTTAG	413
DB	181	TGAATGGGAAAGACCAAGCTGGCACTGCAATGTACTTTACTGGAGCAGATGTTTGT	473
QY	414	TCAAGCTCAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGTTTGT	300
DB	241	TCAACCTCAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTGGAGTATGATGGTTGGT	300
QY	474	ACACGCTGTAATGTGCAACTAAACACAGCTGAAAGAAATTTAA	516
DB	301	GACTGCTGTAATGTGCAACTAAACACAGCTGAAAGAAATTTAA	343
RESULT 13			
US-10-466-894-390			
; Sequence 390, Application US/10466894			
; Publication No. US20040241671A1			
; GENERAL INFORMATION:			
; APPLICANT: Telerman, Adam			
; APPLICANT: Amson, Robert			
; APPLICANT: Tuijnder, Marius			
; APPLICANT: Susini, Laurent			
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,			
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE			
; TITLE OF INVENTION: AND THEIR USE AS MEDICINES			
; FILE REFERENCE: 10918-014-999			
; CURRENT APPLICATION NUMBER: US/10/466,894			
; CURRENT FILING DATE: 2003-07-23			
; PRIOR APPLICATION NUMBER: PCT/FR02/00273			
; PRIOR FILING DATE: 2002-01-23			
; PRIOR APPLICATION NUMBER: FR01/00899			
; PRIOR FILING DATE: 2001-01-23			
; NUMBER OF SEQ ID NOS: 2270			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 390			
; LENGTH: 343			
; TYPE: DNA			
; ORGANISM: Homo Sapiens			
US-10-466-894-390			
Query Match 14.1%; Score 283.8; DB 18; Length 343;			
Best Local Similarity 89.2%; Pred. No. 2.9e-64;			
Matches 306; Conservative 0; Mismatches 37; Indels 0; Gaps 0			
QY	174	TTTACACGAGGAGCGAAACTACTTTGGACCAACCAAGCGAGGACATGGAATGGAAACCCAC	233
DB	1	TTTACACGAGGAGCGGAGGACCACTTTGGACCAACCAAGCGAGGACATGGAATGGAGAGCGC	60
QY	234	AGTGGATGAAATGACCAAAAGCAAGTATTCAATTTTGTATCTTTGGTTAAAAATGTTT	293
DB	61	AGTGGATGACATGACCAAAAGCAAGTATTCAATTTTGTATCTTTGGTTAAAAATGTTT	120
QY	294	ATTTGAAGTGTCTTAACAAAGAATATATTTCTCTGGTGATGTTAAATTTGGTTTGTGCAACA	353
DB	121	GTTCGAAGTGTCTAGCAAAAGCAATAGCTCTCTAGTAAATTTACTTTGGTTCTGTCAGCA	180
QY	354	TGAATGGGAAAGACCAAGCTGGCACTGCAATGTACTTAATTTGGAGGAAGACATTTAG	413
DB	181	TGAATGGGAAAGACCAAGCTGGCACTGCAATGTACTTTACTGGAGCAGATGTTTGT	473
QY	414	TCAAGCTCAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGTTTGT	300
DB	241	TCAACCTCAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTGGAGTATGATGGTTGGT	300
QY	474	ACACGCTGTAATGTGCAACTAAACACAGCTGAAAGAAATTTAA	516

Db 301 GAGTGCCTGTAATGTTCAACTAACACGAGCTGAAAAA 343

RESULT 14

US-10-466-894-433
; Sequence 433, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; PRIOR FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 433
; LENGTH: 403
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 218
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-433

Query Match 14.18; Score 283.8; DB 18; Length 403;
Best Local Similarity 89.0%; Pred. No. 3.2e-64;
Matches 306; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
Qy 173 CTTTACAACGAGGAGCGGAACTACTTTGGACCAAAAGCGAGGACATGGAATGGAAACCA 232
Db 56 CTTTACAACGAGGAGCGGAGACCACTTTGGACCAAAAGCGAGGACATGGAATGGAGAGCG 115
Qy 233 CAGTGGATGAATGACCAAAAGCAAGTATTTCATTTTGTGTTTAAATAATGTT 292
Db 116 CAGTGGATGACATGACCAAAAGCAAGTATTTCATTTTGTGTTTAAATAATGTT 175
Qy 293 TATTTGAAGTGCTTAACACAAAGNATATATTCCTGCTGATGTTAATTTGTCGAAC 352
Db 176 TGTGTAAGTGCTCAGACAAAGAACATAGCTCCTAGTAAATGNTACTTTGTTGTCGAGC 235
Qy 353 ATGAATGGGAAAGAACCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAAGACTTTTA 412
Db 236 ATGATGGGAAAGAACCAAGGCTGGCACTGTCTATGCTGATTTGGAGGCAAGACTTTTA 295
Qy 413 GTCAAGCTCAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTGAGCAGATGTTGG 472
Db 296 GTCAACCTCAAGGAAATGTTGGAGAGGCAAGCTAAATGTTGTTACTGAGTAGATGTTGG 355
Qy 473 TAACAGCTGTAATGTGCAACTAACACCACTGCAAGCTGAAAGAAATTTAA 516
Db 356 TGACTGCTGTAATGTTCAACTAACACCACTGCAAGCTGAAAGAAATTTAA 399

RESULT 15

US-10-466-894-369
; Sequence 369, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent

; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; PRIOR FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 369
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 448,475
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-369

Query Match 14.0%; Score 283.2; DB 18; Length 497;
Best Local Similarity 89.0%; Pred. No. 5.2e-64;
Matches 306; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
Qy 173 CTTTACAACGAGGAGCGGAACTACTTTGGACCAAAAGCGAGGACATGGAATGGAAACCA 232
Db 68 CTTTACAACGAGGAGCGGAGACCACTTTGGACCAAAAGCGAGGACATGGAATGGAGAGCG 127
Qy 233 CAGTGGATGAATGACCAAAAGCAAGTATTTCATTTTGTGTTTAAATAATGTT 292
Db 128 CAGTGGATGACATGACCAAAAGCAAGTATTTCATTTTGTGTTTAAATAATGTT 187
Qy 293 TATTTGAAGTGCTTAACACAAAGNATATATTCCTGCTGATGTTAATTTGTCGAAC 352
Db 188 TGTGTAAGTGCTCAGACAAAGAACATAGCTCCTAGTAAATGTTACTTTGTTGTCGAGC 247
Qy 353 ATGAATGGGAAAGAACCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAAGACTTTTA 412
Db 248 ATGAATGGGAAAGAACCAAGGCTGGCACTGTCTATGCTGATTTGGAGGCAAGACTTTTA 307
Qy 413 GTCAAGCTCAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTGAGCAGATGTTGG 472
Db 308 GTCAACCTCAAGGAAATGTTGGAGAGGCAAGCTAAATGTTGTTACTGAGTAGATGTTGG 367
Qy 473 TAACAGCTGTAATGTGCAACTAACACCACTGCAAGCTGAAAGAAATTTAA 516
Db 368 TGACTGCTGTAATGTTCAACTAACACCACTGCAAGCTGAAAGAAATTTAA 411

Search completed: January 23, 2005, 14:35:13
Job time : 1074 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2005, 23:46:58 ; Search time 6354.75 Seconds
(without alignments)
11577.453 Million cell updates/sec

Title: US-10-069-056-12
Perfect score: 2019
Sequence: 1 atggctggaatgcttactc.....agccgtgaacttgactaa 2019

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gssi:*
9: gb_gssi2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	61.6	3.1	1101	9	CNS0039G
C 2	59.8	3.0	997	9	CNS005TE
C 3	50	2.5	1134	8	CC245800
C 4	50	2.5	1241	9	AG448181
C 5	49.8	2.5	1101	9	CNS0006J
C 6	49.8	2.5	1101	9	CNS0182P
C 7	49.6	2.5	1101	9	CNS017XK
C 8	49.4	2.4	886	9	CNS075DF
C 9	49.2	2.4	913	9	CG772152
C 10	49.2	2.4	930	9	CG770435
C 11	48.4	2.4	1083	9	CNS07A2Y
C 12	47.8	2.4	987	9	CNS00418
C 13	47.6	2.4	1025	9	CNS014J2
C 14	47.2	2.3	1811	9	CG753732
C 15	46.6	2.3	933	9	CNS00CFN
C 16	46.4	2.3	1080	9	CNS00EPP
C 17	46.2	2.3	1758	9	CL509408
C 18	46	2.3	769	9	CNS011WN
C 19	45.8	2.3	963	9	CNS00A4L
C 20	45.8	2.3	1066	6	CD386086
C 21	45.8	2.3	1204	9	CNS016E2
C 22	45.6	2.3	1200	9	CNS016EL
C 23	45.4	2.2	975	6	CD050911
C 24	45.2	2.2	644	8	BH463121

BOHEK14TR

C 25	45.2	2.2	979	8	AZ538989
C 26	45	2.2	768	9	CNS01VSE
C 27	45	2.2	894	9	CNS018BG
C 28	45	2.2	1184	8	B13117
C 29	44.8	2.2	368	2	AW767845
C 30	44.8	2.2	414	4	BJ640957
C 31	44.8	2.2	435	8	BZ665713
C 32	44.8	2.2	530	9	CE641746
C 33	44.8	2.2	581	9	CL553259
C 34	44.8	2.2	1843	9	AG435185
C 35	44.6	2.2	298	4	B1815258
C 36	44.6	2.2	298	4	B1815258
C 37	44.6	2.2	363	5	BQ451492
C 38	44.6	2.2	870	8	AZ693037
C 39	44.6	2.2	1199	8	CC240969
C 40	44.4	2.2	706	9	AG178839
C 41	44.4	2.2	878	9	CNS0187R
C 42	44.4	2.2	938	9	CNS006TJ
C 43	44.4	2.2	1201	5	BX416077
C 44	44.2	2.2	588	1	AJ773389
C 45	44.2	2.2	694	1	AJ778527
C 46	44.2	2.2	1500	9	AG334092

ALIGNMENTS

RESULT 1
CNS0039G/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION
AL063921

VERSION
AL063921.1

KEYWORDS
GSS.

SOURCE
Drosophila melanogaster (fruit fly)

ORGANISM
Drosophila melanogaster

..
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
1 (bases 1 to 1101)

Genoscope.

AUTHORS
Direct Submission

TITLE
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

JOURNAL
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

COMMENT
- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).

The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mamooser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
Location/Qualifiers

1..1101

/organism="Drosophila melanogaster"

/mol_type="genomic DNA"

/db_xref="taxon:7227"

/clone="BACR08K10"

/clone_lib="RPCI-98"

/note="end : TET3"

ORIGIN

Query Match 3.1%; Score 61.6; DB 9; Length 1101;

Best Local Similarity 16.2%; Pred. No. 0.00018;

CNS0182P 1101 bp DNA linear GSS 26-JUL-1999
Drosophila melanogaster genome survey sequence SP6 end of BAC
BACN37D10 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

AL108811
AL108811.1 GI:5629115
GSS.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

FEATURES
source
1. .1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN37D10"
/clone_lib="DrosBAC"
/plasmid="pBelobAC11"
/note="end : SP6"

ORIGIN
Query Match 2.5%; Score 49.8; DB 9; Length 1101;
Best Local Similarity 21.9%; Pred. No. 0.2;
Matches 115; Conservative 179; Mismatches 228; Indels 3; Gaps 1;
207 AACGAGGACATGGATGGAAACACACAGTGGATGAATGACCAAAAGCAAGTATTCAT 266
517 AAGGAAAAAANAANAANAANAANAAGGKGAAGGGRGRBVBGAGDGTATAATAA 576
267 TTTTGAATCTTGGTAAAAAATGTTTATTGAAGTCTTAACACAAAGATATATTC 326
577 AAADTKORDARGAARAAAAAADAARWAARWAARWAARWAARWAARWAARWAAR 636
327 TGGTGTATTTAATTTGTTGTCAACATGAATGGGAAAGACCAAGGCTGGCACTGCCA 386
637 TTCTTAAAAAAGGGGKTKGKAADGDGGAADAADWRGRGRGRRAAARAAAAADAKR 696
387 TGTACTAATTTGGAGGAAGACATTTAGTCAAGCTCAAGGAAATGGTGGAGGCAACT 446
697 AAAAAAANAANAANAANAAGGKRGKGDGDKATKTAAWAKGKDWGTATAMTWD 756
447 AAATGTTTACTGGACGAGATGGTTGTGAACAGCTGTAATGTGCAACTAACACAGCTGA 506
757 ATWADTWAATDADAKRAAAGRRKRDARKTARGGRPARTR---RRAAAGGRRARAGAR 813
507 AAGAATTAATAAGAGAAATAGCAGAAACAATGAGTGGGTTACTCTACTTACTTATAA 566
814 RAAARAAADRDWAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 873
567 GCATAAGCAACCAAAAGACTATACCAAGTGTGTTCTTTTGGAAACATGATGCTTA 626
874 RARARRRRRRRRARRARRAADDTDTKRWADATTDKTTTWTDDDDWDKAKRD 933
627 CTATTTTAACTAAAGAATAAGCACTAGTCCACCAAGACGAGGCTATTTCCT 686
934 RWAAXADGAWKWRDRAADWAATAKDDDDGKDKGGRKGRKGRKGRKGRKGRKGRK 993

CNS017KX 1101 bp DNA linear GSS 26-JUL-1999
Drosophila melanogaster genome survey sequence SP6 end of BAC
BACN37F10 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

AL108171
AL108171.1 GI:5628475
GSS.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

FEATURES
source
1. .1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN37F10"
/clone_lib="DrosBAC"
/plasmid="pBelobAC11"
/note="end : SP6"

ORIGIN
Query Match 2.5%; Score 49.6; DB 9; Length 1101;
Best Local Similarity 21.8%; Pred. No. 0.23;
Matches 95; Conservative 143; Mismatches 198; Indels 0; Gaps 0;
260 TATTCATTTTGAATCTTTGGTTAAAAAATGTTTATTGAAAGTGTTAACACAAAGATA 319
1077 TWKTADKWTWDTKDDGGGGGGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGK 1018
320 TATTTCTCTGTGATGTTAATTTGTTTGTCAACATGAATGGGAAAGACCAAGGCTGC 379
1017 DDTWKTGTGTTKTKTKWKKWARRAAKWRDWRDWDKWDTKWDRWTKRWATRDGR 958
380 ACTGCATCTACTAATTTGGAGAAAGCACTTTAGTCAAGCTCAAGGAAATGGTGGAGAA 439
957 AGRARERDRGRDRGRRRRRRTTRKWRWRRAAWTAWTTTWWGTTKGAWTKRKGK 898
440 GGCAACTAAATGTTTACTGAGCAGATGGTTGGTAAACAGCTGTAATGTCACATAAC 499
897 GAAAAADWGTDDWRWDKRAAAAAAGKGGKGGKGGKGGKGGKGGKGGKGGKGGK 838
500 CAGCTGAAGAATAAAGCAAGAAATAGCAGACATGATGAGTGGTGTACTTACTTCTTA 559
837 KTGRAWAAAGTGTWDAATWAKTKRAAAAAGRAAAAANKDRGTGKGGKGGKGGTDDGK 778
560 CTTATAAGCAATAGCAACCAAAAAAGACTATACCAAGTGTGTTCTTTTGGAAAAATGA 619
777 TWKTTTWTWTAARARARAAAKGWTKTDTGRBARAWDANTTTTTRKATRDGAA 718
620 TTGCTTACTATTTTAACTAAAAAAGAAATAAGACACTAGTCCACCAAGACGAGGCT 679

QY 687 TAGCACTGACTCTGGCTGGAAAACTAACTTTTAAAGAGGCGCA 731
Db 994 DMDKTWTRDWWMTTKMDWDDGGRGRGRWTRKRWGAWWRADA 1038

RESULT 7

CNS017KX/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence SP6 end of BAC
BACN37F10 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

COMMENT

FEATURES
source

ORIGIN

Query Match 2.5%; Score 49.6; DB 9; Length 1101;
Best Local Similarity 21.8%; Pred. No. 0.23;
Matches 95; Conservative 143; Mismatches 198; Indels 0; Gaps 0;

QY 260 TATTCATTTTGAATCTTTGGTTAAAAAATGTTTATTGAAAGTGTTAACACAAAGATA 319
Db 1077 TWKTADKWTWDTKDDGGGGGGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGK 1018

QY 320 TATTTCTCTGTGATGTTAATTTGTTTGTCAACATGAATGGGAAAGACCAAGGCTGC 379
Db 1017 DDTWKTGTGTTKTKWKKWARRAAKWRDWRDWDKWDTKWDRWTKRWATRDGR 958

QY 380 ACTGCATCTACTAATTTGGAGAAAGCACTTTAGTCAAGCTCAAGGAAATGGTGGAGAA 439
Db 957 AGRARERDRGRDRGRRRRRRTTRKWRWRRAAWTAWTTTWWGTTKGAWTKRKGK 898

QY 440 GGCAACTAAATGTTTACTGAGCAGATGGTTGGTAAACAGCTGTAATGTCACATAAC 499
Db 897 GAAAAADWGTDDWRWDKRAAAAAAGKGGKGGKGGKGGKGGKGGKGGKGGKGGK 838

QY 500 CAGCTGAAGAATAAAGCAAGAAATAGCAGACATGATGAGTGGTGTACTTACTTCTTA 559
Db 837 KTGRAWAAAGTGTWDAATWAKTKRAAAAAGRAAAAANKDRGTGKGGKGGKGGTDDGK 778

QY 560 CTTATAAGCAATAGCAACCAAAAAAGACTATACCAAGTGTGTTCTTTTGGAAAAATGA 619
Db 777 TWKTTTWTWTAARARARAAAKGWTKTDTGRBARAWDANTTTTTRKATRDGAA 718

QY 620 TTGCTTACTATTTTAACTAAAAAAGAAATAAGACACTAGTCCACCAAGACGAGGCT 679


```

/strain="GA-2"
/db_xref="taxon:7070"
/clone_lib="Tribolium BAC library"
/notes="Vector: pBACE3.6; Site 1: EcoRI; Site 2: EcoRI;
Library constructed by Exelixis Inc."

ORIGIN
Query Match      2.4%; Score 49.2; DB 9; Length 913;
Best Local Similarity 58.0%; Pred. No. 0.27;
Matches 87; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 23 ATCAAGTTTGGGAGCAACCACTGGTTAAAGGAAAAGTACCAAGGAGTCTTCAT 82
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 296 ATGAATTTTACACAGACAATTAATTACATCTTTAAATGTAATTTGGAAGTATCTGAA 355
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 83 TTGTTTTTAAATGAAAATGTTCAACTGAATGAAAAGATATCGGATGGAATAGTTACA 142
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 356 CTATTTTATAACAAAATCTTCGAAGACAGAAAATAAATACGACTAACTAAA 415
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 143 AAAAGAGCTGCAGGAGGAGAGCTGAAAT 172
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 416 AAATAACTATCGACAACCTCTGTACTAAAAT 445
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
CG770435      930 bp DNA linear GSS 29-OCT-2003
LOCUS      TcB52.2 A06 SP6 Tribolium BAC library Tribolium castaneum genomic,
DEFINITION      genomic survey sequence.
ACCESSION      CG770435
VERSION      CG770435.1 GI:38023614
KEYWORDS      GSS.
SOURCE      Tribolium castaneum (red flour beetle)
ORGANISM      Tribolium castaneum
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
Tenebrionidae; Tribolium.
REFERENCE      1 (bases 1 to 930)
AUTHORS      Savard, J. and Tautz, D.
TITLE      Tribolium castaneum BAC-ends sequencing project
JOURNAL      Unpublished (2003)
COMMENT      Contact: Savard, J.
Abteilung fur Evolutionsgenetik, AG Tautz
Institut fur Genetik, Universitat zu Koln
Weyertal 121, 50931 Koln, Germany
Tel: 49 221 470 6911
Fax: 49 221 470 5975
Email: savard@uni-koeln.de
Class: BAC ends.
FEATURES             source
    1..930
        Location/Qualifiers
            /organism="Tribolium castaneum"
            /mol_type="genomic DNA"
            /strain="GA-2"
            /db_xref="taxon:7070"
            /clone_lib="Tribolium BAC library"
            /notes="Vector: pBACE3.6; Site 1: EcoRI; Site 2: EcoRI;
Library constructed by Exelixis Inc."

ORIGIN
Query Match      2.4%; Score 49.2; DB 9; Length 930;
Best Local Similarity 58.0%; Pred. No. 0.28;
Matches 87; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 23 ATCAAGTTTGGGAGCAACCACTGGTTAAAGGAAAAGTACCAAGGAGTCTTCAT 82
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 288 ATGAATTTTACACAGACAATTAATTACATCTTTAAATGTAATTTGGAAGTATCTGAA 347
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 83 TTGTTTTTAAATGAAAATGTTCAACTGAATGAAAAGATATCGGATGGAATAGTTACA 142
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 348 CTATTTTATAACAAAATCTTCGAAGACAGAAAATAAATACGACTAACTAAA 407
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 143 AAAAGAGCTGCAGGAGGAGAGCTGAAAT 172
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
408 AAATAACTATCGACAACCTCTGTACTAAAAT 437

CNS07A2Y      1063 bp DNA linear GSS 08-JUL-2001
LOCUS      T3 end of clone BC0AA001B01 of library BC0AA from strain CBS 767 of
DEFINITION      Debaryomyces hansenii, genomic survey sequence.
ACCESSION      AL436064
VERSION      AL436064.1 GI:12219477
KEYWORDS      GSS.
SOURCE      Debaryomyces hansenii (anamorph: Candida famata)
ORGANISM      Debaryomyces hansenii
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Debaryomycetes.
REFERENCE      1 (bases 1 to 1063)
AUTHORS      Souciet, J.-L., Aigle, M., Artiguenave, F., Blandin, G.,
Bolotin-Fukuhara, M., Bon, E., Brottier, P., Casaregola, S.,
de-Montigny, J., Dujon, B., Durrens, P., Lepingle, A., Llorente, B.,
Malpertuy, A., Neuveglise, C., Ozier-Kalogeropoulos, O., Potier, S.,
Saurin, W., Tekai, F., Toffano-Nioche, C., Wesolowski-Louvel, M.,
Wincker, P. and Weissenbach, J.
Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
FEBS Lett. 487 (1), 3-12 (2000)
20584711
11152876
REFERENCE      2 (bases 1 to 1063)
AUTHORS      Lepingle, A., Casaregola, S., Neuveglise, C., Bon, E., Nguyen, H.,
Artiguenave, F., Wincker, P. and Gaillardin, C.
Genomic exploration of the hemiascomycetous yeasts: 14.
Debaryomyces hansenii var. hansenii
FEBS Lett. 487 (1), 82-85 (2000)
20584724
11152889
REFERENCE      3 (bases 1 to 1063)
AUTHORS      Genoscope.
TITLE      Direct Submission
JOURNAL      Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex - FRANCE (E-mail :
seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT      This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxi,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.
FEATURES             source
    1..1063
        Location/Qualifiers
            /organism="Debaryomyces hansenii"
            /mol_type="genomic DNA"
            /strain="CBS 767"
            /variety="hansenii"
            /db_xref="taxon:4959"
            /clone="BC0AA001B01"
            /clone_lib="BC0AA"
            /notes="end : T3"

ORIGIN
Query Match      2.4%; Score 48.4; DB 9; Length 1063;
Best Local Similarity 38.2%; Pred. No. 0.46;
Matches 108; Conservative 38; Mismatches 137; Indels 0; Gaps 0;

QY 369 CCAAGCTGTCACATGCGCATCTACTAATTTGGAGGAAGGACTTTAGTCAAGCTCAAGGAA 428
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 661 CCCGSSSSSSSACCGCGTTCGNTTCCTCGSTGSGTTTTTTTTTWWWKWWATWYAT 720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 429 ATCGTGGAGAGGCAACTAAATGTTTACTGGAGAGATGTTGTTAAACAGCTGTAATGT 488
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



```
Db      876 AAGGTGTGATWAGWRAAAATTTTWTTRTAAATGGTGATTAARWAAAAAAGHAKAGA 817
Qy      380 ACTGCCATGTAATAATGGAGGAAAGGACTTTAGTCAAGCTCAAGGAAATGCTGGAGAA 439
Db      816 RKGG---AGAAAAATGKGGTGWRAAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 761
Qy      440 GGCACAACTAAATGTTTACTGGAGCAGATGGTGTGTAAACAGCCTGTAATGTCACACTAACAC 499
Db      760 GTAAAAAATAATGTTGTTAAATAAAATTTTCTGATGTAARAWAAAAATATAAAT 701
Qy      500 CAGCTGAAGAAATTAACCTAAGAGAAATAGCAGACATGAGTGGGTTACTCTACTTA 559
Db      700 AWAYGAATGTAAGAAAAAARAAAAATGTAWATAWATTTKATTTGTAATGWTGTTK 641
Qy      560 CTTTATAAGCAATAGCAAAACCAAAAAAGACTATACCAAGTGTGTTCTTTTGGAAAAATGA 619
Db      640 GTKKAWAAAAAAGAAAAAAGKRWTAWKTARAAAWDGTGKTWKTCTTTTAAWAAAGGR 581
Qy      620 TTGCTTACTATTTTAACTAAAGAAAGAAATAGCAGCTAGTCCACAGAGCGGAGGCT 679
Db      580 GTGKGAAGTTRTDWAAAAAAGAAAAAAGAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAG 521
Qy      680 ATTTTCTTACGAGTCTG--GCTGGAACCTAACTTTTAAAGAGGCGGAGGCGCA 737
Db      520 AGRWAAAAARATATAGTGTGKTGTGTGTAGTAAAGATWAAATGTAAGGTTGTAKWAT 461
Qy      738 TCTAGTGCAGCAACTATACACTGATGACATGCGGCCAGAAACGGTTGAAACACAGTAAC 797
Db      460 ATTAAAAAAGAAAAAAGAAAAAAGKRWTAWKTARAAAWDGTGKTWKTCTTTTAAWAAAG 401
Qy      798 CACTGCGCAGAACTAAGCGCGGCGAATTCACAACTAAAGAAAGAGTCTTCTATTAAAC 857
Db      400 WTRATKKTGKTGDKAARAADAATTTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 341
Qy      858 TACACTTAAGAGCTGGTGATTAAGAGAGTAACCTCACCAGAGGAGTGGATGATGCA 917
Db      340 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAA 281
Qy      918 GCCAGAGCAGTTACATTTGAAATGATGGCTCAACAGGAGGAGGAGGAGGAGGAGGAGG 977
Db      280 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAA 221
Qy      978 GCTAGAGATTTGACACTACTAGCCAGAAACAAAGAGCAATTTGACTTAATTTTGA 1037
Db      220 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAA 161
Qy      1038 AAGAGTGAACAGCAACCACTAACCACTTTTCACTGCTGACACAGAACCTGCAGAA 1097
Db      160 NNGNANNNNAANNNAANNNAANNNAANNNAANNNAANNNAANNNAANNNAANNNA 101
Qy      1098 TTTTGTCTTTTCACTGCTGGAACATATGTTAAAGTTTGCCTGCTATTGCTGCTTTTAA 1157
Db      100 TATTAGTTTNNNNNNNANDANTATWNTANNNNNNNNNNNNNNNNNNNNNNNNNN 41
Qy      1158 CAGACAAGGAGGCAAAAGAAATG 1180
Db      40 NNNNANNNNNCCACCAATTAAG 18
```

```
RESULT 14
CG753732/c
LOCUS
DEFINITION
  CG753732 1811 bp DNA linear GSS 24-OCT-2003
  P048-4-G03-ya Ppa EcoRI BAC Library Pristionchus pacificus genomic,
  genomic survey sequence.
ACCESSION
  CG753732
VERSION
  CG753732.1 GI:37978509
KEYWORDS
  GSS.
SOURCE
  Pristionchus pacificus
  Pristionchus pacificus
  Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
  Neodiplogasteridae; Pristionchus.
REFERENCE
  1 (bases 1 to 1811)
```

```
AUTHORS
  Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
  Bunjer,J., van der Neulen,M. and Sommer,R.J.
TITLE
  An integrated physical and genetic map of the nematode Pristionchus
  pacificus
JOURNAL
  Mol. Genet. Genomics 269 (5), 715-722 (2003)
MEDLINE
  22835951
PUBMED
  12884007
COMMENT
  Contact: Sommer RJ
  Evolutionary Biology
  Max-Planck-Institute for Developmental Biology
  Spemannstr. 37-39, Tuebingen D-72076, Germany
  Tel: 00497071601371
  Fax: 00497071601498
  Email: ralf.sommer@tuebingen.mpg.de
  Class: BAC ends.
FEATURES
  source
  1..1811
  /organism="Pristionchus pacificus"
  /mol_type="genomic DNA"
  /strain="California"
  /db_xref="taxon:54126"
  /clone_lib="Ppa EcoRI BAC Library"
  /notes="The library was generated by a partial digest of
  the genomic DNA with EcoRI and cloning into the BAC
  vector."
ORIGIN
  Query Match 2.3%; Score 47.2; DB 9; Length 1811;
  Best Local Similarity 34.1%; Pred. No. 1.1;
  Matches 352; Conservative 0; Mismatches 677; Indels 2; Gaps 1;
  Qy 36 AGCAACCACTGGTTAAAGAAAAAGTAACACGAGGAGTGTCTCTATGTTTAAAAA 95
  Db 1472 ANAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 1413
  Qy 96 TGAATATGTTCACTCAATCGAATGGAAGATATCGGATGGAATAGTACAAAAAGAGCTGCA 155
  Db 1412 AAAAAAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAAN 1353
  Qy 156 GGAGGAGCAGCTGAAATCTTTTCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 215
  Db 1352 NNANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 1293
  Qy 216 CATGGAATGGGAAACCAACAGTGTGAATGACAAAGAGCAAGTATTCATTTTGTATTC 275
  Db 1292 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1233
  Qy 276 TTTGGTTAAAAATGTTTATTTTGAAGTCTTACACAGAGGAGGAGGAGGAGGAGGAGG 335
  Db 1232 NAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1173
  Qy 336 TAATTTGGTTTGTCAACATGATGGGAAAAAGACC--AAGGCTGGCACTGCCATGTA 393
  Db 1172 AAAAAAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAAN 1113
  Qy 394 ATTGGAGGAAAGGACTTTTGTAGTCAAGGCAATGTTGGAGAGGAGGAGGAGGAGGAGG 453
  Db 1112 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1053
  Qy 454 TACTGGAGCAGATGGTTGTTAAGCAGCTGTAATGTGCACTAATACACCAGCTGAAGGATT 513
  Db 1052 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 993
  Qy 514 AAATAAGAGAAATAGCAGAGGAGCAATGAGTGGGTTACTCTACTTACTTAAAGCATAG 573
  Db 992 ANAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 933
  Qy 574 CAAACCAAAAAAGACTATATACCAAGTGTGTTCTTTTGGAAACATGATGCTTACTATTT 633
  Db 932 NAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 873
  Qy 634 TTAACCTAAAGAAATAGCACTAGTCCCAAGAGAGGAGGAGGAGGAGGAGGAGGAGG 693
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:18:38 ; Search time 99 Seconds
(without alignments)
2435.011 Million cell updates/sec

Title: US-10-069-056-14

Perfect score: 3637

Sequence: 1 MAGNAYSDEVLGATNWLKEK.....RACFGAEPLKDFSEPLNLD 672

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3627	99.7	672	4	AAy72708	Parvoviru
2	3623	99.6	672	4	AAy72702	Parvoviru
3	3620	99.5	672	4	AAy72704	Parvoviru
4	3618	99.5	672	4	AAy72710	Parvoviru
5	3618	99.5	672	4	AAy72706	Parvoviru
6	2376	65.3	690	1	AAp40306	Sequence
7	459.5	12.6	626	2	AAr85384	Barbary d
8	459.5	12.6	626	4	AAy97724	NS1 prote
9	459.5	12.6	626	5	AAE22891	Muscovy d
10	459.5	12.6	626	5	AAE28648	Muscovy d
11	459.5	12.6	626	5	AAE26944	Muscovy d
12	459.5	12.6	626	6	ABU64869	NS1 prote
13	459.5	12.6	626	6	ABU64764	Parvoviru
14	459.5	12.6	626	6	ABR43402	Muscovy d
15	459.5	12.6	626	7	ADi40288	Muscovy d
16	459	12.6	627	4	AAy97722	Rep prote
17	459	12.6	627	5	AAE22889	Barbarie
18	459	12.6	627	5	AAE28646	Barbarie
19	459	12.6	627	5	AAE26942	Barbarie
20	459	12.6	627	6	ABU64867	Rep prote
21	459	12.6	627	6	ABU64762	Parvoviru
22	459	12.6	627	6	ABR43400	Barbarie
23	459	12.6	627	7	ADi40284	Barbarie
24	453.5	12.5	550	3	AAy58164	Adeno ass
25	453.5	12.5	550	5	AAU11409	Adeno-ass

26	453.5	12.5	610	3	AAy58159	Adeno ass
27	453.5	12.5	610	4	AAy97720	Rep prote
28	453.5	12.5	610	5	AAU11404	Adeno-ass
29	453.5	12.5	610	5	AAE22887	Adeno-ass
30	453.5	12.5	610	5	AAE28644	Adeno-ass
31	453.5	12.5	610	5	AAE26940	Adeno ass
32	453.5	12.5	610	6	ABU64865	Rep prote
33	453.5	12.5	610	6	ABU64760	Adeno ass
34	453.5	12.5	610	6	ABR43398	Adeno-ass
35	453.5	12.5	610	7	ADi40280	Adeno-ass
36	453.5	12.5	610	8	ADH58895	Predeterm
37	440.5	12.1	625	7	ABR62760	Adeno ass
38	440.5	12.1	627	4	AAy97723	Rep prote
39	440.5	12.1	627	5	AAE28647	Goose par
40	440.5	12.1	627	5	AAE26943	Goose par
41	440.5	12.1	627	6	ABU64868	Rep prote
42	440.5	12.1	627	6	ABU64763	Parvoviru
43	440.5	12.1	627	6	ABR43401	Goose par
44	440.5	12.1	627	7	ADi40286	Goose par
45	434	11.9	627	5	AAE22890	Goose par

ALIGNMENTS

RESULT 1
AAy72708

ID AAY72708 standard; protein; 672 AA.

XX AAY72708;

XX 06-AUG-2003 (revised)

DT 31-MAY-2001 (first entry)

XX

DE Parvovirus non-structure protein 1 (NS1) variant (T394A).

XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;

KW tumoural disease; gene therapy; mutant; mutein; variant.

XX

OS Parvovirus.

OS Synthetic.

XX

PH Key Location/Qualifiers

FT Misc-difference 394

FT /note= "Wild type Thr substituted with Ala"

XX

PN EP1077260-A1.

XX 21-FEB-2001.

XX 13-AUG-1999; 99EP-00115161.

XX 13-AUG-1999; 99EP-00115161.

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX Nueesch J, Rommelaere J;

XX WPI; 2001-212717/22.

XX N-PSDB; AAD02803.

PT Novel parvovirus non-structure protein variant, useful for treating tumoral diseases, has a shifted equilibrium between DNA replication and transcription activities, and cytotoxic activity.

XX Claim 6; Page 25-27; 4lpp; English.

XX The present sequence is parvovirus non-structure protein 1 (NS1) variant (T394A). The invention relates to the variants of the parvovirus non-structure protein (NS1) having a shifted equilibrium between the DNA replication and transcription activities, and the cytotoxicity activity. These variants are useful as toxins for treating tumoral diseases. The variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-

```
CC 2003 to correct OS field.)
XX
SQ Sequence 672 AA;

Query Match      99.7%; Score 3627; DB 4; Length 672;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 671; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVLGATNWLKESKNOEVFSFVKENNVQNGKDIGNSYKKELQDELKSLQ 60
DB 1 MAGNAYSDEVLGATNWLKESKNOEVFSFVKENNVQNGKDIGNSYKKELQDELKSLQ 60
QY 61 RGAETTWDSQSEDMEWETTVDKQVFIKQVIFDLSLVKKCLFEVLTNKNIFPGDVNMFVQHEW 120
DB 61 RGAETTWDSQSEDMEWETTVDKQVFIKQVIFDLSLVKKCLFEVLTNKNIFPGDVNMFVQHEW 120
QY 121 GKDOGWCHVLIIGKDFSOAQGWRRQLNVYWSRWLVACNVQLTPAERIKLRETAEDN 180
DB 121 GKDOGWCHVLIIGKDFSOAQGWRRQLNVYWSRWLVACNVQLTPAERIKLRETAEDN 180
QY 181 EWTLLTYKHQKQTKDYTKCVLFGNNMIAVYFLTKKISTSPPRDGGYFLSSDSGKTNFL 240
DB 181 EWTLLTYKHQKQTKDYTKCVLFGNNMIAVYFLTKKISTSPPRDGGYFLSSDSGKTNFL 240
QY 241 KEGERHLVSKLYTDDMRPETVETVTTAQTETKGRGRIQTKKEVSIKTTLKELVHKRVTSPE 300
DB 241 KEGERHLVSKLYTDDMRPETVETVTTAQTETKGRGRIQTKKEVSIKTTLKELVHKRVTSPE 300
QY 301 DNMMPQDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFTDLILEKAETSKLTNFSLPD 360
DB 301 DNMMPQDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFTDLILEKAETSKLTNFSLPD 360
QY 361 TRTCRIFAFHGWNVYKCHAIACVLRQGGKRNVLPHGPASTGKSIQAIAQAQVGNVG 420
DB 361 TRTCRIFAFHGWNVYKCHAIACVLRQGGKRNVLPHGPASTGKSIQAIAQAQVGNVG 420
QY 421 CYNAAVNVFPNDCTNKNLIWVEEAGNFGQVNOFKAIACSGQTIRIDQKGGSKQIEPTP 480
DB 421 CYNAAVNVFPNDCTNKNLIWVEEAGNFGQVNOFKAIACSGQTIRIDQKGGSKQIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVK 540
DB 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVK 540
QY 541 NGYQSTMASYCAKWKVPDSENWABPKVPTPINLLGSARSPTTPKSTPLSQNYALTPL 600
DB 541 NGYQSTMASYCAKWKVPDSENWABPKVPTPINLLGSARSPTTPKSTPLSQNYALTPL 600
QY 601 ASDLEDLALPWSPTNTPVAGTAETQNTGEGSKACQDGLSPTWSEIEDLACFGAEP 660
DB 601 ASDLEDLALPWSPTNTPVAGTAETQNTGEGSKACQDGLSPTWSEIEDLACFGAEP 660
QY 661 LKXDFSEPLNLD 672
DB 661 LKXDFSEPLNLD 672
```

RESULT 2

```
AA72702
ID AAY72702 standard; protein; 672 AA.
XX
AC AAY72702;
XX
DT 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
DE Parvovirus wild-type non-structure protein 1 (NS1).
XX
KW NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoural disease; gene therapy.
XX
OS Parvovirus.
XX
```

```
PN EP1077260-A1.
XX
PD 21-FEB-2001.
XX
PF 13-AUG-1999; 99EP-00115161.
XX
PR 13-AUG-1999; 99EP-00115161.
XX
PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
PI Nueesch J, Rommelaere J;
XX
DR WPI; 2001-212717/22.
DR N-PSDB; AAD02797.
XX
PT Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX
PS Disclosure; Fig 1; 41pp; English.
XX
CC The present sequence is a parvovirus wild-type non-structure protein 1
CC (NS1). The present invention relates to the variants of the parvovirus
CC non-structure protein (NS1) having a shifted equilibrium between the DNA
CC replication and transcription activities, and the cytotoxicity activity.
CC These variants are useful as toxins for treating tumoural diseases. The
CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX
SQ Sequence 672 AA;
```

```
Query Match      99.6%; Score 3623; DB 4; Length 672;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 670; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVLGATNWLKESKNOEVFSFVKENNVQNGKDIGNSYKKELQDELKSLQ 60
DB 1 MAGNAYSDEVLGATNWLKESKNOEVFSFVKENNVQNGKDIGNSYKKELQDELKSLQ 60
QY 61 RGAETTWDSQSEDMEWETTVDKQVFIKQVIFDLSLVKKCLFEVLTNKNIFPGDVNMFVQHEW 120
DB 61 RGAETTWDSQSEDMEWETTVDKQVFIKQVIFDLSLVKKCLFEVLTNKNIFPGDVNMFVQHEW 120
QY 121 GKDOGWCHVLIIGKDFSOAQGWRRQLNVYWSRWLVACNVQLTPAERIKLRETAEDN 180
DB 121 GKDOGWCHVLIIGKDFSOAQGWRRQLNVYWSRWLVACNVQLTPAERIKLRETAEDN 180
QY 181 EWTLLTYKHQKQTKDYTKCVLFGNNMIAVYFLTKKISTSPPRDGGYFLSSDSGKTNFL 240
DB 181 EWTLLTYKHQKQTKDYTKCVLFGNNMIAVYFLTKKISTSPPRDGGYFLSSDSGKTNFL 240
QY 241 KEGERHLVSKLYTDDMRPETVETVTTAQTETKGRGRIQTKKEVSIKTTLKELVHKRVTSPE 300
DB 241 KEGERHLVSKLYTDDMRPETVETVTTAQTETKGRGRIQTKKEVSIKTTLKELVHKRVTSPE 300
QY 301 DNMMPQDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFTDLILEKAETSKLTNFSLPD 360
DB 301 DNMMPQDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFTDLILEKAETSKLTNFSLPD 360
QY 361 TRTCRIFAFHGWNVYKCHAIACVLRQGGKRNVLPHGPASTGKSIQAIAQAQVGNVG 420
DB 361 TRTCRIFAFHGWNVYKCHAIACVLRQGGKRNVLPHGPASTGKSIQAIAQAQVGNVG 420
QY 421 CYNAAVNVFPNDCTNKNLIWVEEAGNFGQVNOFKAIACSGQTIRIDQKGGSKQIEPTP 480
DB 421 CYNAAVNVFPNDCTNKNLIWVEEAGNFGQVNOFKAIACSGQTIRIDQKGGSKQIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVK 540
DB 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLVK 540
QY 541 NGYQSTMASYCAKWKVPDSENWABPKVPTPINLLGSARSPTTPKSTPLSQNYALTPL 600
DB 541 NGYQSTMASYCAKWKVPDSENWABPKVPTPINLLGSARSPTTPKSTPLSQNYALTPL 600
```

```
Db 541 NGYSTWASYCAKWKGVDPDSENWAEKVPPTPINLIGSARSPTTTPKSTPLSQNYALTPL 600
QY 601 ASDLEDLALPWSPTNTPVAGTAETQNTGEGSKACQDQGLSPTWSEIEDLRACFGAEP 660
Db 601 ASDLEDLALPWSPTNTPVAGTAETQNTGEGSKACQDQGLSPTWSEIEDLRACFGAEP 660
QY 661 LKKDFSEPLNLD 672
Db 661 LKKDFSEPLNLD 672

RESULT 3
ID AAY72704 standard; protein; 672 AA.
XX AC AAY72704;
XX 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX Parvovirus non-structure protein 1 (NS1) variant (S283A).
XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
XX tumoural disease; gene therapy; mutant; mutein; variant.
XX Parvovirus.
OS Synthetic.
XX Key Location/Qualifiers
XX Misc-difference 283 /note= "Wild type Ser substituted with Ala"
FT FT
XX XX
XX EP1077260-A1.
XX 21-FEB-2001.
XX 13-AUG-1999; 99EP-00115161.
XX 13-AUG-1999; 99EP-00115161.
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX Nueesch J, Rommelaere J;
XX WPI; 2001-212717/22.
XX N-PSDB; AAD02799.
XX Novel parvovirus non-structure protein variant, useful for treating
XX tumoural diseases, has a shifted equilibrium between DNA replication and
XX transcription activities, and cytotoxic activity.
XX Claim 6; Page 14-16; 41pp; English.
XX The present sequence is parvovirus non-structure protein 1 (NS1) variant
XX (S283A). The invention relates to the variants of the parvovirus non-
XX structure protein (NS1) having a shifted equilibrium between the DNA
XX replication and transcription activities, and the cytotoxicity activity.
XX These variants are useful as toxins for treating tumoural diseases. The
XX variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
XX 2003 to correct OS field.)
XX SQ Sequence 672 AA;
Query Match 99.5%; Score 3620; DB 4; Length 672;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 669; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 MAGNAYSDEVILGATNWLKESNQEVFVFVKNEENVQLNGKDIGNWSYKKEQLQDELSLQ 60
Db 1 MAGNAYSDEVILGATNWLKESNQEVFVFVKNEENVQLNGKDIGNWSYKKEQLQDELSLQ 60
QY 61 RGAETTWDOSEDMEWETTVDKQVFI FDSL VKKCLFEVLNTKNI FPGDVNWFVQHEW 120
|||||
```

```
Db 61 RGAETTWDOSEDMEWETTVDKQVFI FDSL VKKCLFEVLNTKNI FPGDVNWFVQHEW 120
QY 121 GKDQGHCHVLIGGKDFSOAQGKWRRLQNLVYWSRWLVTAACNVQLTPAERI KUREIAEDN 180
|||||
Db 121 GKDQGHCHVLIGGKDFSOAQGKWRRLQNLVYWSRWLVTAACNVQLTPAERI KUREIAEDN 180
QY 181 EWTLLTYKHKOTKKDYTKCVLFGNMIAYYELTKKISTSPPRDGGYFLSSDSGSKTNFL 240
|||||
Db 181 EWTLLTYKHKOTKKDYTKCVLFGNMIAYYELTKKISTSPPRDGGYFLSSDSGSKTNFL 240
QY 241 KEGERHLVSKLYTDDMRPETVTTTAQETKRGRIQTKEVSIKTTKLKELVHKRVTSPE 300
|||||
Db 241 KEGERHLVSKLYTDDMRPETVTTTAQETKRGRIQTKEVSIKTTKLKELVHKRVTSPE 300
QY 301 DWMMPQDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILKAEKSKLTNFSLPD 360
|||||
Db 301 DWMMPQDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILKAEKSKLTNFSLPD 360
QY 361 TRTCRIFAFHGWNYVKVCHAI CCVLNRGGKRNVL FHPGASTGKSI IAAQIAQAVGNVG 420
|||||
Db 361 TRTCRIFAFHGWNYVKVCHAI CCVLNRGGKRNVL FHPGASTGKSI IAAQIAQAVGNVG 420
QY 421 CYNAAANVNFPPNDCTNKNLIWVEEAGNFQGVNQFKAI CSQTIRIDQKGGSKQIEPTP 480
|||||
Db 421 CYNAAANVNFPPNDCTNKNLIWVEEAGNFQGVNQFKAI CSQTIRIDQKGGSKQIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPI RDRMLNIHLTHLPDGLVDKNWPMICAWLVK 540
|||||
Db 481 VIMTTNENITVVRIGCEERPEHTQPI RDRMLNIHLTHLPDGLVDKNWPMICAWLVK 540
QY 541 NGYSTWASYCAKWKGVDPDSENWAEKVPPTPINLIGSARSPTTTPKSTPLSQNYALTPL 600
|||||
Db 541 NGYSTWASYCAKWKGVDPDSENWAEKVPPTPINLIGSARSPTTTPKSTPLSQNYALTPL 600
QY 601 ASDLEDLALPWSPTNTPVAGTAETQNTGEGSKACQDQGLSPTWSEIEDLRACFGAEP 660
|||||
Db 601 ASDLEDLALPWSPTNTPVAGTAETQNTGEGSKACQDQGLSPTWSEIEDLRACFGAEP 660
QY 661 LKKDFSEPLNLD 672
Db 661 LKKDFSEPLNLD 672

RESULT 4
ID AAY72710 standard; protein; 672 AA.
XX AC AAY72710;
XX 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX Parvovirus non-structure protein 1 (NS1) variant (T463A).
XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
XX tumoural disease; gene therapy; mutant; mutein; variant.
XX Parvovirus.
OS Synthetic.
XX Key Location/Qualifiers
XX Misc-difference 463 /note= "Wild type Thr substituted with Ala"
FT FT
XX XX
XX EP1077260-A1.
XX 21-FEB-2001.
XX 13-AUG-1999; 99EP-00115161.
XX 13-AUG-1999; 99EP-00115161.
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX PA
```

```
XX Nueesch J, Rommelaere J;
PI WPI; 2001-212717/22.
DR N-PSDB; AAD02805.
XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX Claim 6; Page 30-32; 41pp; English.
XX The present sequence is parvovirus non-structure protein 1 (NS1) variant
CC (T463A). The invention relates to the variants of the parvovirus non-
CC structure protein (NS1) having a shifted equilibrium between the DNA
CC replication and transcription activities, and the cytotoxicity activity.
CC These variants are useful as toxins for treating tumoral diseases. The
CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX SQ Sequence 672 AA;

Query Match 99.5%; Score 3618; DB 4; Length 672;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 669; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MAGNAYSDEVLGATNWLKESNQEVFVFKNENVLQNGKDIGNWSYKKELODELKSLQ 60
Db 1 MAGNAYSDEVLGATNWLKESNQEVFVFKNENVLQNGKDIGNWSYKKELODELKSLQ 60
Qy 61 RGAETTWQSEDMEWETTVDENTKKQVFI FDSL VKKCLFEVLNTKNI FPGDVNMFVQHEW 120
Db 61 RGAETTWQSEDMEWETTVDENTKKQVFI FDSL VKKCLFEVLNTKNI FPGDVNMFVQHEW 120
Qy 121 GKQGWCHVLIGKDSQAGKQWRRQLNYSRWLVTAACNVQLTPAERIKLREIAEDN 180
Db 121 GKQGWCHVLIGKDSQAGKQWRRQLNYSRWLVTAACNVQLTPAERIKLREIAEDN 180
Qy 181 EWTLLTYKHQTKKDYTKCVLFGNMIAYYFLTKKISTSPRDPGGYFLSSDGSWKTNFL 240
Db 181 EWTLLTYKHQTKKDYTKCVLFGNMIAYYFLTKKISTSPRDPGGYFLSSDGSWKTNFL 240

Query Match 99.5%; Score 3618; DB 4; Length 672;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 669; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MAGNAYSDEVLGATNWLKESNQEVFVFKNENVLQNGKDIGNWSYKKELODELKSLQ 60
Db 1 MAGNAYSDEVLGATNWLKESNQEVFVFKNENVLQNGKDIGNWSYKKELODELKSLQ 60
Qy 61 RGAETTWQSEDMEWETTVDENTKKQVFI FDSL VKKCLFEVLNTKNI FPGDVNMFVQHEW 120
Db 61 RGAETTWQSEDMEWETTVDENTKKQVFI FDSL VKKCLFEVLNTKNI FPGDVNMFVQHEW 120
Qy 121 GKQGWCHVLIGKDSQAGKQWRRQLNYSRWLVTAACNVQLTPAERIKLREIAEDN 180
Db 121 GKQGWCHVLIGKDSQAGKQWRRQLNYSRWLVTAACNVQLTPAERIKLREIAEDN 180
Qy 181 EWTLLTYKHQTKKDYTKCVLFGNMIAYYFLTKKISTSPRDPGGYFLSSDGSWKTNFL 240
Db 181 EWTLLTYKHQTKKDYTKCVLFGNMIAYYFLTKKISTSPRDPGGYFLSSDGSWKTNFL 240
```

```
RESULT 5
AAY72706
ID AAY72706 standard; protein; 672 AA.
XX
XX AAY72706;
AC AC
XX 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
XX Parvovirus non-structure protein 1 (NS1) variant (T363A).
DE NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
XX tumoral disease; gene therapy; mutant; mutein; variant.
XX Parvovirus.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Misc-difference 363
FT /note= "Wild type Thr substituted with Ala"
FT
XX EP1077260-Al.
PN 21-FEB-2001.
XX
XX 13-AUG-1999; 99EP-00115161.
PF
XX 13-AUG-1999; 99EP-00115161.
PR
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
XX Nueesch J, Rommelaere J;
XX WPI; 2001-212717/22.
DR N-PSDB; AAD02801.
XX
XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX
XX Claim 6; Page 19-21; 41pp; English.
XX The present sequence is parvovirus non-structure protein 1 (NS1) variant
CC (T363A). The invention relates to the variants of the parvovirus non-
CC structure protein (NS1) having a shifted equilibrium between the DNA
CC replication and transcription activities, and the cytotoxicity activity.
CC These variants are useful as toxins for treating tumoral diseases. The
CC variant DNAs are useful as vectors for gene therapy. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX
XX SQ Sequence 672 AA;

Query Match 99.5%; Score 3618; DB 4; Length 672;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 669; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MAGNAYSDEVLGATNWLKESNQEVFVFKNENVLQNGKDIGNWSYKKELODELKSLQ 60
Db 1 MAGNAYSDEVLGATNWLKESNQEVFVFKNENVLQNGKDIGNWSYKKELODELKSLQ 60
Qy 61 RGAETTWQSEDMEWETTVDENTKKQVFI FDSL VKKCLFEVLNTKNI FPGDVNMFVQHEW 120
Db 61 RGAETTWQSEDMEWETTVDENTKKQVFI FDSL VKKCLFEVLNTKNI FPGDVNMFVQHEW 120
Qy 121 GKQGWCHVLIGKDSQAGKQWRRQLNYSRWLVTAACNVQLTPAERIKLREIAEDN 180
Db 121 GKQGWCHVLIGKDSQAGKQWRRQLNYSRWLVTAACNVQLTPAERIKLREIAEDN 180
Qy 181 EWTLLTYKHQTKKDYTKCVLFGNMIAYYFLTKKISTSPRDPGGYFLSSDGSWKTNFL 240
Db 181 EWTLLTYKHQTKKDYTKCVLFGNMIAYYFLTKKISTSPRDPGGYFLSSDGSWKTNFL 240
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:27:50 ; Search time 25 Seconds
(without alignments)
1782.628 Million cell updates/sec

Title: US-10-069-056-14
Perfect score: 3637
Sequence: 1 MAGNAYSDEVLGATNWLKKE.....RACFGAEPLKDFSEPLNLD 672

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 56318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	432.5	11.9	536	4 US-09-532-594B-10	Sequence 10, Appl
2	432.5	11.9	623	4 US-09-532-594B-2	Sequence 2, Appl
3	428.5	11.8	546	4 US-09-807-802A-7	Sequence 7, Appl
4	428.5	11.8	623	4 US-09-807-802A-2	Sequence 2, Appl
5	428.5	11.8	623	4 US-09-807-802A-5	Sequence 5, Appl
6	428.5	11.7	312	4 US-09-532-594B-8	Sequence 8, Appl
7	425.5	11.7	399	4 US-09-532-594B-9	Sequence 9, Appl
8	425.5	11.6	322	4 US-09-807-802A-11	Sequence 11, Appl
9	420.5	11.6	399	4 US-09-807-802A-9	Sequence 9, Appl
10	420.5	11.6	399	4 US-09-807-802A-3	Sequence 3, Appl
11	139	3.8	276	3 US-09-171-461-3	Sequence 3, Appl
12	139	3.8	276	3 US-09-970-711-3	Sequence 3, Appl
13	124.5	3.4	647	3 US-09-300-909-17	Sequence 17, Appl
14	114	3.1	564	3 US-09-134-001C-2897	Sequence 897, Ap
15	113	3.1	589	4 US-09-107-532A-6086	Sequence 6086, Ap
16	104	2.9	1170	4 US-09-638-524A-2	Sequence 2, Appl
17	101	2.8	715	4 US-09-248-796A-20602	Sequence 20602, A
18	101	2.8	1120	4 US-09-792-024-95	Sequence 95, Appl
19	100.5	2.8	2409	6 US-09-08-2	Patent No. 5180808
20	100	2.7	481	4 US-09-252-991A-30005	Sequence 30005, A
21	100	2.7	605	4 US-09-107-532A-6077	Sequence 6077, Ap
22	100	2.7	1503	4 US-09-677-046A-6	Sequence 6, Appl
23	100	2.7	1509	4 US-09-677-046A-2	Sequence 2, Appl
24	99.5	2.7	657	3 US-09-300-909-16	Sequence 16, Appl
25	99	2.7	919	4 US-09-248-796A-19408	Sequence 19408, A
26	97.5	2.7	556	4 US-09-248-796A-18763	Sequence 18763, A
27	97.5	2.7	3248	1 US-08-353-700-1	Sequence 1, Appl

28	97.5	2.7	3248	5 PCT-US95-16216-1	Sequence 1, Appl
29	96.5	2.7	557	4 US-09-489-039A-11233	Sequence 11233, A
30	96	2.6	291	4 US-09-107-532A-6390	Sequence 6390, Ap
31	95.5	2.6	1060	4 US-09-248-796A-16624	Sequence 16624, A
32	95.5	2.6	3898	3 US-08-750-717-2	Sequence 2, Appl
33	95	2.6	590	4 US-09-134-000C-4755	Sequence 4755, Ap
34	94	2.6	370	4 US-09-583-110-3405	Sequence 3405, Ap
35	94	2.6	471	4 US-09-134-000C-6228	Sequence 6228, Ap
36	94	2.6	605	2 US-08-472-666-1	Sequence 1, Appl
37	94	2.6	605	5 PCT-US96-07615-1	Sequence 1, Appl
38	94	2.6	1115	2 US-08-568-459A-2	Sequence 2, Appl
39	94	2.6	1115	2 US-08-487-826B-2	Sequence 2, Appl
40	94	2.6	1115	3 US-09-210-288-2	Sequence 2, Appl
41	94	2.6	1115	6 US-09-347-6	Patent No. 5198347
42	94	2.6	1307	1 US-08-395-246C-2	Sequence 2, Appl
43	94	2.6	1343	4 US-09-270-767-31792	Sequence 31792, A
44	94	2.6	1343	4 US-09-270-767-47009	Sequence 47009, A
45	94	2.6	1477	4 US-09-206-942-71	Sequence 71, Appl

ALIGNMENTS

RESULT 1
US-09-532-594B-10
; Sequence 10, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AA4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; OTHER INFORMATION: synthetic construct
; NAME/KEY: misc feature
; OTHER INFORMATION: AA4 Rep protein 68
US-09-532-594B-10

Query Match	11.9%	Score	432.5	DB	4	Length	536
Best Local Similarity	27.5%	Pred. No.	8.4e-36				
Matches	156	Conservative	75	Mismatches	233	Indels	103
Gaps	19						
Qy	63	AETTQDSDEMEWETTVDEMTKKQVFDFSLVKCLFEVLNTKNIIPGDVNMVFOHSEWGK	122				
Db	31	AEKEWELPPDSMDLNLIE--QAPLTVAEKQLQRELFVEMRVRVK--APEALFFVQPEKG-	85				
Qy	123	DQGMHCHVLIG-----GKDFSOAQGKWMRQLNVYMSRWLVLTACNVQLTPAERIKL	173				
Db	86	DSYFHLHLIVETGVKSMVGVRYVSQIKER-----LVTRIYRGVEP-----	126				
Qy	174	RETAEDNEWTLTYKHQ--TKKDYTKCVLFGNMIAYFLTKKKKISTSPRGGYFLSS	231				
Db	127	---QLPNPFAVTKTRNGAGGNKVDDC-----YIPNYLLPK-----TQP-----	163				
Qy	232	DSGWK--TNF-----LKEGERHLVSKLYTDDMRPVTETVTTTAQETKRGH-----	275				
Db	164	ELQWAMTWMDQYISACLNTAERKELVAQHL-----THVSQTQEQNKQNPNSD	212				
Qy	276	---IQTKKEVSIKTLKELVHKVTSPEDEMMQPOSYIEMMAQPGGNNLLKNTLEICTL	332				
Db	213	APVIRKTSARYMELVGLVDRGITSEKQWIQEDQASYSIFNAASRSRQKKAALDNASK	272				

QY 393 NAVLPHCPASTGKSIITAAQIAQAVNGVCCYNAANVPFPNDCTKNLIIWVEEAGNFGQOV 452
Db 328 NTIWLFGPATTKNTIAEATAHAYVPFGVCYNWTNENFPFNDKVDKMWIWEEGKMTAKV 387
QY 453 NQKALCSGOTIRIDOKGSKGKQIEPTPIVMTNENITVVRIGCEERPEHTQIRDRMLN 512
Db 388 ESAKAILGSKVRVDQCKSSAQIDPTPIVITNTNMCVIDGNSTTFEHOQLQDMMFK 447
QY 513 IHLTHLPGDFGLVDKNEWPMICAWLVKNGYQSTMASYCAKWG--KVPDWSE--NWAEPKV 569
Db 448 FELTKRLEHDFGKVKQEVKDFERWASDHVTEVTFEYVRKGGARKRPADNDADISEPKR 507
QY 570 PTINLLGSARSPTTPKSTPLSQNYA 596
Db 508 ACP-----SVAQPSDAPV--DYA 527

RESULT 4
US-09-807-802A-7
; Sequence 7, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNVN.031USA
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US/09/807, 802A
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 7
; LENGTH: 546
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-7

Query Match 11.8%; Score 428.5; DB 4; Length 546;
Best Local Similarity 26.9%; Pred. No. 2.3e-35;
Matches 156; Conservative 63; Mismatches 211; Indels 151; Gaps 20;

QY 63 AETTWDQSEDMEWE-----TTVDMTKKQVFIQDLSLVKCKLFEVLNTKNIFPGDVNW 114
Db 31 AEKEWELPPSDMDLNLIEOAPLTVAEKLDQDFLVQWRRVSKA-----PEALF 78
QY 115 FVQHEGKQGWCHVLIQ-----GKDFSAQAGKWMR-----ROLNVVWSRLVATCN 162
Db 79 FVQFEKGESY-FHLHLVETTVGKSMVLGRFLSQIRDKLVQTIYRGIEPTLPNFWAVT-- 135
QY 163 VQITPAERIKLREIAEDNEWVTLTYKHQTKDYKCVLFGNMIAYYFTKKKISTSP 222
Db 136 -----KTRNGAGGKNKVDECY-----IPNYLLPK-----TOP- 163
QY 223 RDGGYFLSSDSGWK--TNF-----LKEGERHLSKLYTDDMRPETVTTTAAQETK 272
Db 164 -----ELQAWNTMEEYISACNLAEKRLVAQHL-----THVSQTOEQN 203
QY 273 RGRIOTKKEYSI---KTTLKE-----LVHKRVTSPEDMMMQPDSEYIEMMAQPGGNLL 323
Db 204 KENLNPNSDAPVIRSKTSARYMELVGLVDRGITSEKQWIOEQDQASVISFNAASNSRQI 263
QY 324 KNTLEICTLTARTKTAFLILKEAETSKLTFNLSLPTTRCFRIF-----A 368
Db 264 KAALDNAGKIMALTAKSAPDVLVGPAPPA-----DICTNRYIRILELNGYEPAYAGSV 315
QY 369 FHGWNVYKVCCHAICCVNLNROGGKRNALVPHGPASTGKSIITAAQIAQAVNGVCCYNAANVN 428
Db 316 FLGW-----AQKRFGRKNTIWLFGPATTKNTIAEATAHAYVPFGVCNWTNEN 363

QY 429 FPENDCTKNLIIWVEEAGNFGQOVNOFKALCSGOTIRIDOKGSKGKQIEPTPIVMTNEN 488
Db 364 FPFNDCTVDKMWIWEEGKMTAKYVESAKAILGSKVRVDQCKSSAQIDPTPIVITSNTN 423
QY 489 ITVVRIGCEERPEHTQIRDRMLNLIHTHLPGDFGLVDKNEWPMICAWLVKNGYQSTMA 548
Db 424 MCAVIDGNSTTFEHOQLQDMMFKFELTRLEHDFGKVKQEVKEPRWAQDHVTEVAHE 483
QY 549 SYCAKWG-----KVPDWSNENMAEPKVPPTPINLLGSARSPTTP 585
Db 484 FYVRKGGANKRPAPDDADK-SEPKRACP-----SVADPSTS 518

RESULT 5
US-09-807-802A-2
; Sequence 2, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNVN.031USA
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US/09/807, 802A
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 623
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-2

Query Match 11.8%; Score 428.5; DB 4; Length 623;
Best Local Similarity 26.9%; Pred. No. 2.9e-35;
Matches 156; Conservative 63; Mismatches 211; Indels 151; Gaps 20;

QY 63 AETTWDQSEDMEWE-----TTVDMTKKQVFIQDLSLVKCKLFEVLNTKNIFPGDVNW 114
Db 31 AEKEWELPPSDMDLNLIEOAPLTVAEKLDQDFLVQWRRVSKA-----PEALF 78
QY 115 FVQHEGKQGWCHVLIQ-----GKDFSAQAGKWMR-----ROLNVVWSRLVATCN 162
Db 79 FVQFEKGESY-FHLHLVETTVGKSMVLGRFLSQIRDKLVQTIYRGIEPTLPNFWAVT-- 135
QY 163 VQITPAERIKLREIAEDNEWVTLTYKHQTKDYKCVLFGNMIAYYFTKKKISTSP 222
Db 136 -----KTRNGAGGKNKVDECY-----IPNYLLPK-----TOP- 163
QY 223 RDGGYFLSSDSGWK--TNF-----LKEGERHLSKLYTDDMRPETVTTTAAQETK 272
Db 164 -----ELQAWNTMEEYISACNLAEKRLVAQHL-----THVSQTOEQN 203
QY 273 RGRIOTKKEYSI---KTTLKE-----LVHKRVTSPEDMMMQPDSEYIEMMAQPGGNLL 323
Db 204 KENLNPNSDAPVIRSKTSARYMELVGLVDRGITSEKQWIOEQDQASVISFNAASNSRQI 263
QY 324 KNTLEICTLTARTKTAFLILKEAETSKLTFNLSLPTTRCFRIF-----A 368
Db 264 KAALDNAGKIMALTAKSAPDVLVGPAPPA-----DICTNRYIRILELNGYEPAYAGSV 315
QY 369 FHGWNVYKVCCHAICCVNLNROGGKRNALVPHGPASTGKSIITAAQIAQAVNGVCCYNAANVN 428
Db 316 FLGW-----AQKRFGRKNTIWLFGPATTKNTIAEATAHAYVPFGVCNWTNEN 363
QY 429 FPENDCTKNLIIWVEEAGNFGQOVNOFKALCSGOTIRIDOKGSKGKQIEPTPIVMTNEN 488
Db 364 FPFNDCTVDKMWIWEEGKMTAKYVESAKAILGSKVRVDQCKSSAQIDPTPIVITSNTN 423

QY 489 ITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDGLVDKNEWPMICAWLVKNGYQSTMA 548
Db 424 MCAVIDGNSTTFEHOQPLQDRMPKFLTRLEHDFGKVTQKEVKEPRFRAQDHVTEVAHE 483
QY 549 SYCAKMG-----KVPDWSNWAEPKVPPTPINLLGSARSPTFT 585
Db 484 FYVRKGANKRPAPDDADK-SEPKRACP-----SVADPSTS 518
RESULT 6
US-09-807-802A-5
; Sequence 5, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807.802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 623
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-5

Query Match 11.8%; Score 428.5; DB 4; Length 623;
Best Local Similarity 26.9%; Pred. No. 2.9e-35;
Matches 156; Conservative 63; Mismatches 211; Indels 151; Gaps 20;
QY 63 AETTWDQSEDMWE-----TTVDGMTKKQVFIQFDSLVKKCLFEVLNTKNIFPGGVNW 114
Db 31 AKEWELPPSDMDLNLIEQAPLTVAEKLRDFLQVQWRYSKA-----PEALF 78
QY 115 FVQHEWKGQGWCHVLIG-----GKDFSAQAGKWR-----RQNVVWSRWLVTAQN 162
Db 79 FVQFEKESY-FHLHLVETTVGKSMVLGRFLSQIRKLVQTIYRGIEPTLPNFAVT-- 135
QY 163 VQTPAERIKLRETAEDNEWVTLTYKHQTKDXYKCVLFGNMIAVYFLTKKISTSP 222
Db 136 -----KTRNAGGGNKVDECY-----IPNLLPK-----TOP- 163
QY 223 RDGGYFLSSDSGWK-TNF-----LKEGERHLVSKLYTDDMRPETVTTVTTAQETK 272
Db 164 -----ELQAWTWNMEYISACLNLAERKRLVAQHL-----THVSQTQEQN 203
QY 273 RGRIOYKVESI---KTLTK-----LVKRVTSBEDNMOPDSYIENMAOPGGENLL 323
Db 204 KENLNPSADPVRISKTSARYMELVGLVDGITSEKQWIQEDQASYSIFNAAASNSRQI 263
QY 324 KNTLEICTLTARTKTAFLDILEKAETSKLTNFSLPDTRTCRIF-----A 368
Db 264 KAALDNAGKIMALKSPADVLVGPAP-----DIKNRIYRILELNGYEPVAGSV 315
QY 369 FHGNNYKVKCHAI CCVLNROGGKRNVLPHGPASTGKSIITAOIAQAQVGNVGCYNAVN 428
Db 316 FLGW-----AQKRFGRKNTIWLFGPATTKTNIAEAIAHAVPYGCNVNTNEN 363
QY 429 FPNDCNKNLIWVEEAGNGQOVNQFKAICSGQTRIDQKGSQIEPTPVTMTNEN 488
Db 364 FPNDCVDRKMWIWEEGKMTAKVVESAKAILGSKVRVDQCKSSAQIDPTPIVITSNTN 423
QY 489 ITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDGLVDKNEWPMICAWLVKNGYQSTMA 548
Db 424 MCAVIDGNSTTFEHOQPLQDRMPKFLTRLEHDFGKVTQKEVKEPRFRAQDHVTEVAHE 483

QY 549 SYCAKMG-----KVPDWSNWAEPKVPPTPINLLGSARSPTFT 585
Db 484 FYVRKGANKRPAPDDADK-SEPKRACP-----SVADPSTS 518
RESULT 7
US-09-532-594B-8
; Sequence 8, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532.594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 Rep protein 40
US-09-532-594B-8

Query Match 11.7%; Score 425.5; DB 4; Length 312;
Best Local Similarity 35.3%; Pred. No. 1.7e-35;
Matches 109; Conservative 45; Mismatches 140; Indels 15; Gaps 5;
QY 291 LVHKRVTSBEDNMOPDSYIENMAOPGGENLNKNTLEICTLTARTKTAFLDILEKAET 350
Db 7 LVDRGITSEKQWIQEDQASYSIFNAAASNSRQIKAALDNASKIMSLTKTAPDVLVQGNPP 66
QY 351 SKLTNFSLPDTRTCRIFAFHGNNYKVKCHAI CCVLNROGGKRNVLPHGPASTGKSIITAO 410
Db 67 EDIS-----NRIYRILELNGYDIPQYAAVSFLGWAOKFKGKNTIWLFGPATTKTNIAE 121
QY 411 ALAQAVNGVGCYNAANVNFNDCTNKNLIWVEEAGNGQOVNQFKAICSGQTRIDQKG 470
Db 122 AIAHAPVFGYCVNNTNENFPNDCVDKMWIWEEGKMTAKVVESAKAILGSKVRVDQKC 181
QY 471 KGSQIEPTPVTMTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDGLVDKNE 530
Db 182 KSSAQIDPTPIVITSNTNCAVIDGNSTTFEHOQPLQDRMPKFLTRLEHDFGKVTQKE 241
QY 531 WPMICAWLVKNGYQSTMASYCAKMG-----KVPDWSNWAEPKVPPTPINLLGSARSPTTPK 587
Db 242 VKDFFRWASDHVTEVTHFYVRKGANKRPAPDNADISEPKRACP-----SVNQPSIDA 296
QY 588 STPLSQNYA 596
Db 297 EAPV--DVA 303

RESULT 8
US-09-532-594B-9
; Sequence 9, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532.594B
; CURRENT FILING DATE: 2000-03-22


```
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 Rep protein 52
US-09-532-594B-9

Query Match      11.6%; Score 425.5; DB 4; Length 399;
Best Local Similarity 35.3%; Pred. No. 2.7e-35;
Matches 109; Conservative 45; Mismatches 140; Indels 15; Gaps 5;

Qy 291 LVHKRVTSBEDMMQPDSDYIEMMAQPGGNNLLKNTLEICTLTARTKTAFDLILEKAET 350
Db 7 LVDRGITSEKQWIOEDQASYSISFNAASNSRQIKAAALDNAGKIMALTAKSAPDVLVGPAPP 66

Qy 351 SKLTNFSLPDTRTCRIF-----AFHGNVYKVKCHAIACCVLNROGGKENV 395
Db 67 A-----DIKTNRIYRILELNGYBPAYAGSVFLGW-----AKRFGKRNNTI 106

Qy 396 LFHGPASTGKSIIAQAIQAQAVGNVGCYNAANVPFNDCTNKNLIWVEEAGNFGQQVNOF 455
Db 107 WLFPGATTGKTNTAEIAHAHVPYGCNVNTNENFPFNDCCVCKWVWEEGKMTAKVESA 166

Qy 456 KAICSGQTIRIDQKSGKQIEPTPVIMTNNENITVVRIGCEPERPEHTQPIRDRMLNIHL 515
Db 167 KAILGSKVRVDQCKSSAQIDPTPIVITSNTMCAVIDGNSITTFEHOQPLQDRMPKFEL 226

Qy 516 THHLPDGLVDKNEPWCAMLVKNGYQSTMASYCAKMG-----KVPDWSNNAEPKVP 571
Db 227 TRLEHDFGKVTQKVEKFFRWAQDHVTEVAHBFYVRKGGANKRPAPDDADK-SEPKRAC 285

Qy 572 PINLLGSARSPPFTT 585
Db 286 P-----SVADPSTS 294

RESULT 10
US-09-807-802A-9
; Sequence 9, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; FILE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 9
; LENGTH: 399
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-9

Query Match      11.6%; Score 420.5; DB 4; Length 399;
Best Local Similarity 34.7%; Pred. No. 8.9e-35;
Matches 109; Conservative 37; Mismatches 123; Indels 45; Gaps 6;

Qy 291 LVHKRVTSBEDMMQPDSDYIEMMAQPGGNNLLKNTLEICTLTARTKTAFDLILEKAET 350
Db 7 LVDRGITSEKQWIOEDQASYSISFNAASNSRQIKAAALDNAGKIMALTAKSAPDVLVGPAPP 66

Qy 351 SKLTNFSLPDTRTCRIF-----AFHGNVYKVKCHAIACCVLNROGGKENV 395
Db 67 A-----DIKTNRIYRILELNGYBPAYAGSVFLGW-----AKRFGKRNNTI 106

Qy 396 LFHGPASTGKSIIAQAIQAQAVGNVGCYNAANVPFNDCTNKNLIWVEEAGNFGQQVNOF 455
Db 107 WLFPGATTGKTNTAEIAHAHVPYGCNVNTNENFPFNDCCVCKWVWEEGKMTAKVESA 166

Qy 456 KAICSGQTIRIDQKSGKQIEPTPVIMTNNENITVVRIGCEPERPEHTQPIRDRMLNIHL 515
Db 167 KAILGSKVRVDQCKSSAQIDPTPIVITSNTMCAVIDGNSITTFEHOQPLQDRMPKFEL 226

Qy 516 THHLPDGLVDKNEPWCAMLVKNGYQSTMASYCAKMG-----KVPDWSNNAEPKVP 571
Db 227 TRLEHDFGKVTQKVEKFFRWAQDHVTEVAHBFYVRKGGANKRPAPDDADK-SEPKRAC 285

; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 Rep protein 52
US-09-532-594B-9

Query Match      11.7%; Score 425.5; DB 4; Length 399;
Best Local Similarity 35.3%; Pred. No. 2.7e-35;
Matches 109; Conservative 45; Mismatches 140; Indels 15; Gaps 5;

Qy 291 LVHKRVTSBEDMMQPDSDYIEMMAQPGGNNLLKNTLEICTLTARTKTAFDLILEKAET 350
Db 7 LVDRGITSEKQWIOEDQASYSISFNAASNSRQIKAAALDNAGKIMALTAKSAPDVLVGPAPP 66

Qy 351 SKLTNFSLPDTRTCRIFAFHGNVYKVKCHAIACCVLNROGGKENVAFHGPASTGKSIIAQ 410
Db 67 EDLSS-----NRIYRILELNGYBPQYASVFLGWAKKFGKRNITWLFPGATTGKTNTAE 121

Qy 411 AIAQAVGNVGCYNAANVPFNDCTNKNLIWVEEAGNFGQQVNOFKAICSGQTIRIDQK 470
Db 122 AIAHAVPFYGCNVNTNENFPFNDCCVCKWVWEEGKMTAKVESA KAILGSKVRVDQK 181

Qy 471 KGSQKQTEPTVIMTNNENITVVRIGCEPERPEHTQPIRDRMLNIHLTHHLPDGLVDKNE 530
Db 182 KSSAQIDPTPIVITSNTMCAVIDGNSITTFEHOQPLQDRMPKFELTKRLEHDFGKVTQK 241

Qy 531 WPMICAWLVKNGYQSTMASYCAKMG--KVPDWSNNAEPKVPINLLGSARSPPFTPK 587
Db 242 VKDFFRWASHDVTEVTHEFVVRKGGANKRPAPDADISEPKRACP-----SVAQPSTSDA 296

Qy 588 STPLSQNYA 596
Db 297 EAPV--DYA 303

RESULT 9
US-09-807-802A-11
; Sequence 11, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; FILE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 11
; LENGTH: 322
; TYPE: PRT
; ORGANISM: AAV-1
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (943)..(944)
; OTHER INFORMATION: minor splice site
US-09-807-802A-11

Query Match      11.6%; Score 420.5; DB 4; Length 322;
Best Local Similarity 34.7%; Pred. No. 6e-35;
Matches 109; Conservative 37; Mismatches 123; Indels 45; Gaps 6;
```

```
QY 572 PINLGSARSPTT 585
Db 286 P-----SVADPSTS 294

RESULT 11
US-09-171-461-3
; Sequence 3, Application US/09171461
; Patent No. 6335016
; GENERAL INFORMATION:
; APPLICANT: Baker, Adam
; APPLICANT: Cotten, Matthew
; APPLICANT: Chiocca, Susanna
; APPLICANT: Kurzbauer, Robert
; APPLICANT: Schaffner, Gotthold
; TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus
; FILE REFERENCE: 0652.1800000
; CURRENT APPLICATION NUMBER: US/09/171.461
; CURRENT FILING DATE: 1999-01-12
; EARLIER APPLICATION NUMBER: PCT/EP97/01944
; EARLIER FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 276
; TYPE: PRT
; ORGANISM: CELO Virus
; FEATURE:
; OTHER INFORMATION: Position: 1999..2829 /note=ORF2
US-09-171-461-3

Query Match 3.8%; Score 139; DB 3; Length 276;
Best Local Similarity 23.7%; Pred. No. 1.4e-05;
Matches 63; Conservative 41; Mismatches 134; Indels 28; Gaps 9;

QY 286 TTLKELVHKRVTSPEDEMMMQPDSY--IEMMAQPGGENLLKNTLEICTTLTARTKTAFDL 343
Db 10 TLVHALIDRGVWSREQWQVDPAQYQFYHRSKQGFK--VRHILRDVIRHMCWSTRLDDY 67
QY 344 ILEKAETSKLTNFSLPD-----TTCRIFAFHGNVYKVCCHAICCVLNROGGKGNVLFHG 399
Db 68 MSGASTPS-----PDDVLRNPLYQLLCCNGYNPAVVGTTALIRWAGHQ--SNRNTVWIRG 119
QY 400 PASTGKSIIAQAIAQAVGNVGCYNAANVFPNDCTNKNLIWVEEAGNFQGVNQFKAIC 459
Db 120 TPMSGAPYLAQAIAYCSPVGSVDWKNPNPFCGCDPSLVFWWDGGYVYDCCVGLVKQVF 179
QY 460 SGQTRIDQKG---KGSQKIEPTPVIMTNNITVVRIGCEE--RPEHTQPIRDRMLNTH 514
Db 180 RGEHVILPPEGLRGNPCSELFTPTVLMYSQADICMTRLRSGELSAEHAVALRDCMYLIR 239
QY 515 LTHHLPGDF---GLVDKNEWPMICAW 537
Db 240 LTE-----DFDCAGGISCADVQKQFVAW 261

RESULT 12
US-09-171-461-3
; Sequence 3, Application US/09970711
; Patent No. 6773709
; GENERAL INFORMATION:
; APPLICANT: Baker, Adam
; APPLICANT: Cotten, Matthew
; APPLICANT: Chiocca, Susanna
; APPLICANT: Kurzbauer, Robert
; APPLICANT: Schaffner, Gotthold
; TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus
; FILE REFERENCE: 0652.1800001
; CURRENT APPLICATION NUMBER: US/09/970.711
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/171.461
; PRIOR FILING DATE: 1999-01-12

Query Match 3.4%; Score 124.5; DB 3; Length 647;
Best Local Similarity 19.7%; Pred. No. 0.0021;
Matches 110; Conservative 78; Mismatches 172; Indels 197; Gaps 26;

QY 39 GKDIGNSYKKELODELK-SLQRAETTWQSE-----DMWEETT-VDEMYYKQVFIIDS 92
Db 100 GKVGGRNT-RGTLOEISLNVSTQATVYSPVDSGYGNMEVETAEVEVT----- 149

; PRIOR APPLICATION NUMBER: PCT/EP97/01944
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 276
; TYPE: PRT
; ORGANISM: CELO Virus
; FEATURE:
; OTHER INFORMATION: Position: 1999..2829 /note=ORF2
US-09-970-711-3

Query Match 3.8%; Score 139; DB 4; Length 276;
Best Local Similarity 23.7%; Pred. No. 1.4e-05;
Matches 63; Conservative 41; Mismatches 134; Indels 28; Gaps 9;

QY 286 TTLKELVHKRVTSPEDEMMMQPDSY--IEMMAQPGGENLLKNTLEICTTLTARTKTAFDL 343
Db 10 TLVHALIDRGVWSREQWQVDPAQYQFYHRSKQGFK--VRHILRDVIRHMCWSTRLDDY 67
QY 344 ILEKAETSKLTNFSLPD-----TTCRIFAFHGNVYKVCCHAICCVLNROGGKGNVLFHG 399
Db 68 MSGASTPS-----PDDVLRNPLYQLLCCNGYNPAVVGTTALIRWAGHQ--SNRNTVWIRG 119
QY 400 PASTGKSIIAQAIAQAVGNVGCYNAANVFPNDCTNKNLIWVEEAGNFQGVNQFKAIC 459
Db 120 TPMSGAPYLAQAIAYCSPVGSVDWKNPNPFCGCDPSLVFWWDGGYVYDCCVGLVKQVF 179
QY 460 SGQTRIDQKG---KGSQKIEPTPVIMTNNITVVRIGCEE--RPEHTQPIRDRMLNTH 514
Db 180 RGEHVILPPEGLRGNPCSELFTPTVLMYSQADICMTRLRSGELSAEHAVALRDCMYLIR 239
QY 515 LTHHLPGDF---GLVDKNEWPMICAW 537
Db 240 LTE-----DFDCAGGISCADVQKQFVAW 261

RESULT 13
US-09-300-909-17
; Sequence 17, Application US/09300909
; Patent No. 6306580
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PREPARATION OF HUMAN PAPILLOMAVIRUS B1 HAVING
; TITLE OF INVENTION: HELICASE ACTIVITY AND METHOD THEREFOR
; NUMBER OF SEQUENCES: 27
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/300,909
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/083,942
; FILING DATE: 01-MAY-1998
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 647 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-300-909-17

Query Match 3.4%; Score 124.5; DB 3; Length 647;
Best Local Similarity 19.7%; Pred. No. 0.0021;
Matches 110; Conservative 78; Mismatches 172; Indels 197; Gaps 26;

QY 39 GKDIGNSYKKELODELK-SLQRAETTWQSE-----DMWEETT-VDEMYYKQVFIIDS 92
Db 100 GKVGGRNT-RGTLOEISLNVSTQATVYSPVDSGYGNMEVETAEVEVT----- 149
```

QY 93 LVKKCLFEVLNTKNIIPFGDVNVFVQHEWGKQDGHCHVLIGGKDFSOAQGKWRRLNVY 152
Db 150 -----VATNTNGDAEGEGHGSVRBE-----CSSVDSASIDSENQPK----- 185
QY 153 WSRWLVTACNVQVTPAERIKLRIADNEWVTLTYKHKT-----KKDYTKC 200
Db 186 -----SPTAQIKLL-LOSNKKAAMLT-OFKETYGLSFTDLVRTFKSKTKC 230
QY 201 VLFGNMIAYFLTKKISTSPRGGYFLSSDSGKTNFLKEGERHLVSK--LYTODMRP 258
Db 231 T----DWVAALFGVHPTIA-----EGFKTLINKYALYTH----- 260
QY 259 ETVETTVTTAETKRGRIOT-----KKEVSIKTKLKVHVRVTSPEMMQPDYSY 310
Db 261 -----IOSLDTKQGVLLILMLIRYTCGKRVTVGKGLSTLLH-----VPESCMLEPPPKL 309
QY 311 IEMMA-----OPGENLLKNTLE-----ICTLTAR-----TKTAFDL--ILEKAETSKLT-- 354
Db 310 RSPVAALYWRITGISNISVVGTGPWIRQLTVIQHIGDSDVFLSDMVQWAFDNEYTDE 369
QY 355 -----NFSLPDTRTCRIFAFHGN-----YKVCCHAIC-----CVL 385
Db 370 SDIAFYVAMLDCNSNAAFLKSNCOAKYVKDCATWCKHYKRAQKQMSMSQWIKFCRSK 429
QY 386 NRQGG-----KRNVALPHGPASTGKSIIAQIAQ 414
Db 430 CDEGGWRPIVQFLRYQGIFFISFLCALKEFLKGTGPKKNCIVYGPANTGKSHFCMSLMH 489
QY 415 AV-GNVGCYNAANVNPFDCTKNLIWVEAGN--FGQVQVQFKALCSGQTRIRIDOKGK 471
Db 490 FLOQTIVISYVNSTSHFWLEPLADAKLMLDDATGTGCSYFNDYMRNALDGYAISLDRKYK 549
QY 472 GSKQIEPTVPVIMTTNEN 488
Db 550 SLLQMKCPPLITSNTN 566

RESULT 14
US-09-134-001C-2897
; Sequence 2897, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134, 001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064, 964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055, 779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2897
; LENGTH: 564
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2897

Query Match 3.1%; Score 114; DB 3; Length 564;
Best Local Similarity 21.4%; Pred. No. 0.021;
Matches 81; Conservative 63; Mismatches 128; Indels 106; Gaps 21;

QY 188 YKHQTKDYTKCVLFGNMITYAVFLTKKISTSPRGGYFLSSDSGW-KTNFLKEGE-- 244
Db 87 YSYFGTKKPNNDVTLVEMMIKVFHI--KRFNSE-----LFFVKANGWQKIN----GDEL 134
QY 245 RHLVSKL---YTDMPRETVEVTTTAQETKRGRIQTKKEVSIKTKLKVHVRVTSPE 301
Db 135 QGLISKMIQVLLVDYKPSL-----STLKNVVDGLQKSTDV 169
QY 302 WMMMQPDSYIEMMAQPGENL-----LKNLTLEICTLT-LARTKTAFLILEKAE--- 349

Db 170 BELVENEHVGIC-----GENMFDLNTFQVVKNSIDIFPKTRMLNSLSTNDVITDKIPPYF 224
QY 350 ---TSKLTNFSPLPDTTRCTRIFAFHGNVYKVCCHAICCVLNRQGGKKNVAVLPHGPASTGKS 406
Db 225 KQYMLQLANYD---DDLQYFLFQ-----HT-AVLLTADTKYRRGLILYGAKNGKS 271
QY 407 ITAQIAQAVGNVGCYNAANVNPFDN-----CTNKLIIWVEAGNFGQO---VNOQ 454
Db 272 VVIELVKSPF---YSKDIVSKPLNELEGRFDEKESLIDKSLMASHEIGOSKIQEIKVND 326
QY 455 FKAICSGQTRIDOKGSKQJE---PTPVIMTTNENITVVRIGCERPEHTOPIRDRML 511
Db 327 FKGLSVESMHRDKGK--TQVEVILDLKLIFSTNAINF-----PPEHAKALERRIN 377
QY 512 NIHLTHHL-PGDGFLVVDK 528
Db 378 IIPCEYVVEKADTSLIDK 395

RESULT 15
US-09-107-532A-6086
; Sequence 6086, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107, 532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085, 598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6086:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 589 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (B) LOCATION 1...589
; SEQUENCE DESCRIPTION: SEQ ID NO: 6086:
US-09-107-532A-6086

Query Match 3.1%; Score 113; DB 4; Length 589;
Best Local Similarity 19.8%; Pred. No. 0.029;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:41:01 ; Search time 90.6667 Seconds
(without alignments)
2677.792 Million cell updates/sec

Title: US-10-069-056-14

Perfect score: 3637

Sequence: 1 MAGNAYSDEVLGATNWLKKEK.....RACFGAEPLKDFSEPLNLD 672

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

- Published Applications AA:*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
 - 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
 - 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
 - 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
 - 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	459.5	12.6	626	9	US-09-792-630-29
2	459.5	12.6	626	10	US-09-953-351-29
3	459.5	12.6	626	13	US-10-080-376-29
4	459.5	12.6	626	14	US-10-082-671-35
5	459.5	12.6	626	14	US-10-097-100-29
6	459.5	12.6	626	14	US-10-023-208-29
7	459	12.6	627	9	US-09-792-630-25
8	459	12.6	627	10	US-09-953-351-25
9	459	12.6	627	13	US-10-080-376-25
10	459	12.6	627	14	US-10-082-671-31
11	459	12.6	627	14	US-10-097-100-25
12	459	12.6	627	14	US-10-023-208-25
13	453.5	12.5	610	9	US-09-792-630-21

14	453.5	12.5	610	10	US-09-953-351-21	Sequence 21, Appl
15	453.5	12.5	610	13	US-10-080-376-21	Sequence 21, Appl
16	453.5	12.5	610	14	US-10-082-671-27	Sequence 27, Appl
17	453.5	12.5	610	14	US-10-097-100-21	Sequence 21, Appl
18	453.5	12.5	610	14	US-10-023-208-21	Sequence 21, Appl
19	453.5	12.5	610	14	US-10-375-192-7	Sequence 7, Appl
20	440.5	12.1	625	14	US-10-423-704A-3	Sequence 3, Appl
21	440.5	12.1	627	9	US-09-792-630-27	Sequence 27, Appl
22	440.5	12.1	627	10	US-09-953-351-27	Sequence 27, Appl
23	440.5	12.1	627	13	US-10-080-376-27	Sequence 27, Appl
24	440.5	12.1	627	14	US-10-082-671-33	Sequence 33, Appl
25	440.5	12.1	627	14	US-10-097-100-27	Sequence 27, Appl
26	440.5	12.1	627	14	US-10-023-208-27	Sequence 27, Appl
27	432.5	11.9	537	10	US-09-254-747-10	Sequence 10, Appl
28	432.5	11.9	623	9	US-09-792-630-5	Sequence 5, Appl
29	432.5	11.9	623	10	US-09-953-351-5	Sequence 5, Appl
30	432.5	11.9	623	10	US-09-254-747-2	Sequence 2, Appl
31	432.5	11.9	623	10	US-09-254-747-11	Sequence 11, Appl
32	432.5	11.9	623	13	US-10-080-376-5	Sequence 5, Appl
33	432.5	11.9	623	14	US-10-082-671-11	Sequence 11, Appl
34	432.5	11.9	623	14	US-10-097-100-5	Sequence 5, Appl
35	432.5	11.9	623	14	US-10-023-208-5	Sequence 5, Appl
36	432.5	11.9	623	14	US-10-375-192-5	Sequence 5, Appl
37	431.5	11.9	623	9	US-09-792-630-13	Sequence 13, Appl
38	431.5	11.9	623	10	US-09-953-351-13	Sequence 13, Appl
39	431.5	11.9	623	13	US-10-080-376-13	Sequence 13, Appl
40	431.5	11.9	623	14	US-10-082-671-19	Sequence 19, Appl
41	431.5	11.9	623	14	US-10-097-100-13	Sequence 13, Appl
42	431.5	11.9	623	14	US-10-023-208-13	Sequence 13, Appl
43	431.5	11.9	623	14	US-10-375-192-2	Sequence 2, Appl
44	428.5	11.8	546	15	US-10-696-261-7	Sequence 7, Appl
45	428.5	11.8	546	15	US-10-696-282-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1

US-09-792-630-29
; Sequence 29, Application US/09792630
; Patent No. US20020168640A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Dahiyat, Bassil I.
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES
; FILE REFERENCE: A-70295/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/792,630
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Muscovy duck parvovirus
US-09-792-630-29

Query Match	12.6%	Score	459.5	DB	9	Length	626
Best Local Similarity	25.3%	Pred. No.	6.3e-31				
Matches	165	Conservative	82	Mismatches	263	Indels	141
Gaps	22						
QY	75	WETT-----VDEMTKKQVFIFDSLKKCLFEVLTKNIF-----PGDVNWNVQHEWGK	122				
Db	45	WEPTGIWNMEHVNLPMTVLADKI-----KNIFIQWNNQFNQDETDFDFQLEEGS	93				
QY	123	DQGWCHVLIGKDFSOAGKWWRRQLNVYWSRWLTACNVQLTPAERIK---LREIAD	179				
Db	94	EY-IHLHVCPCGECRSFVLGRY-----MSQIKDSLRLDYEG	129				
QY	180	N-----EWTLTYKHQTKKDYTKCVLFQNNMIAYVFLTKKISTSPRRDGGYFLSSDSG	234				
Db	130	KQVKIPDFWSIYTKKGGQNKVTAA-----YILHLIPIKKQ-----PELOWAFTNPLF	179				
QY	235	WKTNFKEGEHLVSKLYTDDMRPETVETVTTTAQTKRGRIQTKKEVSIKTKTLKELVHK	294				

```
Db 180 TAAALCLQKROELLDAFOESENNAVVEDQASTAAPL-----ISNRAAKYNSNLVDWLIEM 235
Qy 295 RVTSPEDEMMQDPDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFTADLILEKAETSKLT 354
Db 236 GITSEKQWLTKENKESYRSFOATSSNNRQVKAALENARAEMLLTKTATDYLIGKDPVLDIT 295
Qy 355 NFSLPDRTRCIRIFAFHGWNVYKCHAI CCVLNRQGGKRNALVPHGPASTGKSIIAQIAIAQ 414
Db 296 K-----NRIYQILKLNYPQYVGVSLCGWVKREFNKRNAIWLGPATTKTNIAEIAH 350
Qy 415 AVGNVCYNAANVPFNDCTNKNLIWVEEAGNFQOVNQFKAI CSQTIRIDQKGGSK 474
Db 351 AVPFYGCNVNTNENFPFNDVCDKMLIWMEEGKNTKRVESAKAILGGSVAVRVDKCKGVS 410
Qy 475 QIEPTPVMTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNE-----530
Db 411 CIEPTPVIIISNTDMCMIVDGNSTTMEHRIPLERMFQIVLSHKGNGFKISKKEVKEF 470
Qy 531 --W-----PMICAWLVKNGYQSTWASYCAKWKGVDPDSEN-----WAEPK-----568
Db 471 FKWANDNLVPVSEFKVPTNEQTKLTE-----PVERANEPSEPPKIWAPPTRELEEI 524
Qy 569 -----VPTPINLLGSARSPTTPKSTPLSQNYALTPLASDLEDLALPEWSTPNT 617
Db 525 LRASPELFASVAPLP-----SSPDTSPKRRKTRGEYQVRCAMHSL-DNSMNVFECLEC 576
Qy 618 PVAGTAEQTNTGAGSKAC-QDQQLSPTWSEIEEDLRACFGAEPLKDPSE 667
Db 577 ERANFPFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 612
```

RESULT 2

```
US-09-953-351-29
; Sequence 29, Application US/09953351
; Publication No. US20030036643A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Melander, Christian
; APPLICANT: Liu, Hong-Xiang
; APPLICANT: Jin, Cheng He
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FUSION I
; FILE REFERENCE: A-70814/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/953,351
; CURRENT FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/232,960
; PRIOR FILING DATE: 2000-09-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Muscovy duck parvovirus
US-09-953-351-29
```

```
Query Match 12.6%; Score 459.5; DB 10; Length 626;
Best Local Similarity 25.3%; Pred. No. 6.3e-31;
Matches 165; Conservative 82; Mismatches 263; Indels 141; Gaps 22;

Qy 75 WETT-----VDEMTKKQVIFDPSLVKKCLFEVLNTKNIF-----PGDVNWFVQHEWGK 122
Db 45 WEPTGIWMEHVNLPMTLADKI-----KNIFIQRMNQFNQDETDFPFQLEEGS 93

Qy 123 DQGWCHVLIGGKDFSAQAGKWWRRQLNVYWSRWLTACNVQLTPAERIK---LREIAED 179
Db 94 EY-IHLHAVCPGECRSFVLGRY-----MSQIKDSILRDVYEG 129

Qy 180 N-----EWVTLTYKHQTKDYKCVLFGNMIAYFLTKKISTSPRGGYFLSDSG 234
Db 130 KQVKIPDWFISITTKRGQNKVTAA-----YILHYLIPKKQ-----PELOWAFTNMPLE 179
Qy 235 WKTNFKEGEHLVSKLYTDDMRPETVTTVTAQETKRGRIOTKKEVSIKTKLKVHVK 294
Db 180 TAAALCLQKROELLDAFOESENNAVVEDQASTAAPL-----ISNRAAKYNSNLVDWLIEM 235
Qy 295 RVTSPEDEMMQDPDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFTADLILEKAETSKLT 354
```

```
Db 180 TAAALCLQKROELLDAFOESENNAVVEDQASTAAPL-----ISNRAAKYNSNLVDWLIEM 235
Qy 295 RVTSPEDEMMQDPDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFTADLILEKAETSKLT 354
Db 236 GITSEKQWLTKENKESYRSFOATSSNNRQVKAALENARAEMLLTKTATDYLIGKDPVLDIT 295
Qy 355 NFSLPDRTRCIRIFAFHGWNVYKCHAI CCVLNRQGGKRNALVPHGPASTGKSIIAQIAIAQ 414
Db 296 K-----NRIYQILKLNYPQYVGVSLCGWVKREFNKRNAIWLGPATTKTNIAEIAH 350
Qy 415 AVGNVCYNAANVPFNDCTNKNLIWVEEAGNFQOVNQFKAI CSQTIRIDQKGGSK 474
Db 351 AVPFYGCNVNTNENFPFNDVCDKMLIWMEEGKNTKRVESAKAILGGSVAVRVDKCKGVS 410
Qy 475 QIEPTPVMTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNE-----530
Db 411 CIEPTPVIIISNTDMCMIVDGNSTTMEHRIPLERMFQIVLSHKGNGFKISKKEVKEF 470
Qy 531 --W-----PMICAWLVKNGYQSTWASYCAKWKGVDPDSEN-----WAEPK-----568
Db 471 FKWANDNLVPVSEFKVPTNEQTKLTE-----PVERANEPSEPPKIWAPPTRELEEI 524
Qy 569 -----VPTPINLLGSARSPTTPKSTPLSQNYALTPLASDLEDLALPEWSTPNT 617
Db 525 LRASPELFASVAPLP-----SSPDTSPKRRKTRGEYQVRCAMHSL-DNSMNVFECLEC 576
Qy 618 PVAGTAEQTNTGAGSKAC-QDQQLSPTWSEIEEDLRACFGAEPLKDPSE 667
Db 577 ERANFPFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 612
```

RESULT 3

```
US-10-080-376-29
; Sequence 29, Application US/10080376
; Publication No. US20020172968A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Dahiyat, Bassil I.
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES
; FILE REFERENCE: A-70295-2/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/080,376
; CURRENT FILING DATE: 2000-02-19
; PRIOR APPLICATION NUMBER: US 09/792,630
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Muscovy duck parvovirus
US-10-080-376-29
```

```
Query Match 12.6%; Score 459.5; DB 13; Length 626;
Best Local Similarity 25.3%; Pred. No. 6.3e-31;
Matches 165; Conservative 82; Mismatches 263; Indels 141; Gaps 22;

Qy 75 WETT-----VDEMTKKQVIFDPSLVKKCLFEVLNTKNIF-----PGDVNWFVQHEWGK 122
Db 45 WEPTGIWMEHVNLPMTLADKI-----KNIFIQRMNQFNQDETDFPFQLEEGS 93

Qy 123 DQGWCHVLIGGKDFSAQAGKWWRRQLNVYWSRWLTACNVQLTPAERIK---LREIAED 179
Db 94 EY-IHLHAVCPGECRSFVLGRY-----MSQIKDSILRDVYEG 129

Qy 180 N-----EWVTLTYKHQTKDYKCVLFGNMIAYFLTKKISTSPRGGYFLSDSG 234
Db 130 KQVKIPDWFISITTKRGQNKVTAA-----YILHYLIPKKQ-----PELOWAFTNMPLE 179
Qy 235 WKTNFKEGEHLVSKLYTDDMRPETVTTVTAQETKRGRIOTKKEVSIKTKLKVHVK 294
Db 180 TAAALCLQKROELLDAFOESENNAVVEDQASTAAPL-----ISNRAAKYNSNLVDWLIEM 235
Qy 295 RVTSPEDEMMQDPDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFTADLILEKAETSKLT 354
```



```

Db      236   GITSEKQWLTENKESYRSFQATSSNNROVKAALENARAEMLLTKTATDYLIKGDPVLDTIT 295
Qy      355   NFSLPDTRTCRIFAFAHGWNYVKKVCHAI CCVLNRQGGKENAVLFGHPASTGKSIIIAQAATAQ 414
Db      296   K-----NRIYOILKLNNYNPOVGSVLGCWGVKEFKENA IWL YGPATTGKTNI AEATAH 350
Qy      415   AVGNVGCVNAANVPFPNDCTNKNLIWVEEAGNFQQO VNFKAICSGQTIRIDOKKGSK 474
Db      351   AVPFYGCWNVTNENFPFNDCVDCKMLIMWEEGKMTNKVVESA KAILGGSRAVRDQCKGVS 410
Qy      475   QIEBTPVIMTNENITTVRIGCEBERPEHTQPIRDRLMILHLTHLPGDFGLVDKNE-----530
Db      411   CIEPTPVIIITSNTDMCMVDGNSSTMTEHRIPLEERMFOI VLSHKLEGNFGKISKEXREF 470
Qy      531   -W-----PMICAWLVKNGYQSTWASCAKWGKVPDMSEN-----WAEPK-----568
Db      471   FKWANDNLVPVSEFKVPTNEQTKLTE-----PVPERANESEP PKTWAPTREELBEI 524
Qy      569   -----VPTPINILGARSFPFTPKSTPLSQNYALTPLASDL ELDLAL EPWSTPT 617
Db      525   LRASPFLFASVAPLP-----SSPDTSPKRKKTGRGEQVRCAMHS L-DNSMNVFECLEC 576
Qy      618   PVACTAETONTGREGAGKAC-QDGQLSPTWSEIEEDLRACF GAEBLKGFDPSE 667
Db      577   ERANFPFBQSLSGE--NFCNQHG----WYD-----C A FCNELKDDMWNE 612

RESULT 4
US-10-082-671-35
; Sequence 35, Application US/10082671
; Publication No. US20030049647A1
; GENERAL INFORMATION:
; APPLICANT: DAHIYAT, BASSIL
; APPLICANT: LI, MIN
; TITLE OF INVENTION: USE OF NUCLEIC ACID LIBRARIES TO CREATE TOXICOLOGICAL
; TITLE OF INVENTION: PROFILES
; FILE REFERENCE: XEN/001
; CURRENT APPLICATION NUMBER: US/10/082,671
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/270,781
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 626
; TYPE: PRt
; ORGANISM: Muscovy duck parvovirus
US-10-082-671-35
```

```

RESULT 5
US-10-097-100-29
; Sequence 29, Application US/10097100
; Publication No. US20030068649A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Melander, Christian
; APPLICANT: Liu, Hong-Xiang
; APPLICANT: Jin, Cheng He
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF
; FILE REFERENCE: A-70814/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/097,100
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/953,351
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/232,960
; PRIOR FILING DATE: 2000-09-14
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 29
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Muscovy duck parvovirus
; US-10-097-100-29

```

Db 236 GITSEKOWLTENKESYRSFQATSSNNRQVKAALENARAEMLLTKTATDYLGKDPVLDIT 295
Qy 355 NFSLPDTRTCRIFAFPHGWNVYKVCCHAICCVLNROGGKRNALVPHGPASTGKSIIAQAIAQ 414
Db 296 K-----NRIYQILKLNYPQVYVGSVLCGWVKEFNKRNALWLYGPATTKGTNIAEIAH 350
Qy 415 AVGNVCYNAANVPFNDCTNKNLWVEEAGNFQGVNQPKAICSGQTIRIDQKGGSK 474
Db 351 AVPFYGCNVMTNENFPFNDCTNKNLWVEEAGNFQGVNQPKAICSGQTIRIDQKGGSK 410
Qy 475 QIEPTVIMTNTENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNE----- 530
Db 411 CIEPTVILTSNTDMCMVDGNSSTTWEHRIPLERMFQIVLSHKLGNFGKISKEVKEF 470
Qy 531 --W-----PMICAWLVKNGYQSTMASYCAKMGKVPDWSN-----WAEPK----- 568
Db 471 FKWANDNLVPVSEFKVPTNEQTKLE-----PVPERANEPSEPKIWAAPTREBLEEI 524
Qy 569 -----VPTPINLGSARSPTTPKSTPLSONYALTPLASDLELALPFWSTPNT 617
Db 525 LRASPELFASVAPLP-----SSPDTSPKRTKTRGEYQVRCAMHSL-DNSMNVFECLEC 576
Qy 618 PVAGTAEONTGEGSKAC-QDGQLSPTWSEIEBEDLRACFAGBPLKDDPSE 667
Db 577 ERANFPFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 612

RESULT 6

US-10-023-208-29
; Sequence 29, Application US/10023208
; Publication No. US20030124537A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Liu, Yuan-Ching
; TITLE OF INVENTION: PROCARYOTIC LIBRARIES AND USES
; FILE REFERENCE: A-70174-1/RET/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/023,208
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/256,163
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 29
; TYPE: PRT
; ORGANISM: Muscovy duck parvovirus
US-10-023-208-29

Query Match 12.6%; Score 459.5; DB 14; Length 626;
Best Local Similarity 25.3%; Pred. No. 6.3e-31;
Matches 165; Conservative 82; Mismatches 263; Indels 141; Gaps 22;
Qy 75 WETT-----VDEMTKKQVFIFDLSLVKKCLFEVLTNKIF-----PGDVNWFVQHEWGK 122
Db 45 WEPTGIWNHVNLPWVTLADKI-----KNIFIQWVNFQNDQETDFFQLEEGS 93
Qy 123 DQGWCHVLIGKDFSOAQGKWRRLQNVVSRWLVTACNVQVLTAPAEIRK---LREIAD 179
Db 94 EY-IHLHVCPCGECRSFVLGRY-----MSQIKDSILRDVYEG 129
Qy 180 N-----EWVTLTYKHQTKDYTKCVLFGNMAYFYLTKKISTSPPRDGGYFLSSDSG 234
Db 130 KOVKIPDWFISITTKRGQGNKVTAA-----YILHVLIPKKQ-----PELQWAFNTMPLF 179
Qy 235 WKTNFKEGEHLVSKLYTDDMRPETVTTVTAQETKRGRIQTKKEVSIKTKLKVHK 294
Db 180 TAAALCLOKQELLDAFQSEMNNAVQEDQASTAAPL-----ISNRAKNYSNLVDWLIEM 235
Qy 295 RVTSPEDEMMQPDSDYIEMMAQPGENLKNLTLEICTLTARTKTAFDILILEKAETSKLT 354
Db 236 GITSEKOWLTENKESYRSFQATSSNNRQVKAALENARAEMLLTKTATDYLGKDPVLDIT 295
Qy 355 NFSLPDTRTCRIFAFPHGWNVYKVCCHAICCVLNROGGKRNALVPHGPASTGKSIIAQAIAQ 414

Db 296 K-----NRIYQILKLNYPQVYVGSVLCGWVKEFNKRNALWLYGPATTKGTNIAEIAH 350
Qy 415 AVGNVCYNAANVPFNDCTNKNLWVEEAGNFQGVNQPKAICSGQTIRIDQKGGSK 474
Db 351 AVPFYGCNVMTNENFPFNDCTNKNLWVEEAGNFQGVNQPKAICSGQTIRIDQKGGSK 410
Qy 475 QIEPTVIMTNTENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNE----- 530
Db 411 CIEPTVILTSNTDMCMVDGNSSTTWEHRIPLERMFQIVLSHKLGNFGKISKEVKEF 470
Qy 531 --W-----PMICAWLVKNGYQSTMASYCAKMGKVPDWSN-----WAEPK----- 568
Db 471 FKWANDNLVPVSEFKVPTNEQTKLE-----PVPERANEPSEPKIWAAPTREBLEEI 524
Qy 569 -----VPTPINLGSARSPTTPKSTPLSONYALTPLASDLELALPFWSTPNT 617
Db 525 LRASPELFASVAPLP-----SSPDTSPKRTKTRGEYQVRCAMHSL-DNSMNVFECLEC 576
Qy 618 PVAGTAEONTGEGSKAC-QDGQLSPTWSEIEBEDLRACFAGBPLKDDPSE 667
Db 577 ERANFPFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 612

RESULT 7

US-09-792-630-25
; Sequence 25, Application US/09792630
; Patent No. US20020168640A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Dahiyat, Bassil I.
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES
; FILE REFERENCE: A-70295/RET/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/792,630
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 25
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Barbarie duck parvovirus
US-09-792-630-25

Query Match 12.6%; Score 459; DB 9; Length 627;
Best Local Similarity 26.0%; Pred. No. 6.9e-31;
Matches 169; Conservative 78; Mismatches 264; Indels 140; Gaps 23;
Qy 75 WETT-----VDEMTKKQVFIFDLSLVKKCLFEVLTNKIF-----PGDVNWFVQHEWGK 122
Db 45 WEPTGIWNHVNLPWVTLADKI-----KNIFIQWVNFQNDQETDFFQLEEGS 93
Qy 123 DQGWCHVLIGKDFSOAQGKWRRLQNVVSRWLVTACNVQVLTAPAEIRK---LREIAD 179
Db 94 EY-IHLHCCI-----AGNVRSFVLGRYS-----QIKDSILRDVYEG 130
Qy 180 N-----EWVTLTYKHQTKDYTKCVLFGNMAYFYLTKKISTSPPRDGGYFLSSDSG 234
Db 131 KOVKIPDWFISITTKRGQGNKVTAA-----YILHVLIPKKQ-----PELQWAFNTMPLF 180
Qy 235 WKTNFKEGEHLVSKLYTDDMRPETVTTVTAQETKRGRIQTKKEVSIKTKLKVHK 294
Db 181 TAAALCLOKQELLDAFQSEMNNAVQEDQASTAAPL-----ISNRAKNYSNLVDWLIEM 236
Qy 295 RVTSPEDEMMQPDSDYIEMMAQPGENLKNLTLEICTLTARTKTAFDILILEKAETSKLT 354
Db 237 GITSEKOWLTENKESYRSFQATSSNNRQVKAALENARAEMLLTKTATDYLGKDPVLDIT 296
Qy 355 NFSLPDTRTCRIFAFPHGWNVYKVCCHAICCVLNROGGKRNALVPHGPASTGKSIIAQAIAQ 414
Db 297 K-----NRIYQILKLNYPQVYVGSVLCGWVKEFNKRNALWLYGPATTKGTNIAEIAH 351
Qy 415 AVGNVCYNAANVPFNDCTNKNLWVEEAGNFQGVNQPKAICSGQTIRIDQKGGSK 474

Db 352 AVPFYGCWNTNENFPFNDVCDKMLIWWBEGKMTNKVESAAILGGSVAVRDQCKGVS 411
QY 475 QIEPTPVIMTNNITVVRIGCEERPEHTQPIIDRMLNIHLTHLPDGLVDKNE---- 530
Db 412 CIEPTVITNTDMCMIVDGNSTTMEHRIPLEERMFQIVLSHKLGEFGFKISKEVKEF 471
QY 531 --W-----PMICAWLVKNGYQSTMASYCAKWKGVDPWSN-----WAEPK----- 568
Db 472 FKWANDNLVFWVSSEFKVPTNEQTKLTE-----PVERANEPSEPPKIWAPPTREBLEEI 525
QY 569 -----VPTPINLGSARSPTTPKSTPLSQNYALTPLASDLEDLALSPWSTPNT 617
Db 526 LRASPELFASVAPLP-----SSPDTSPKKTGRGEYQVRCAMHSL-DNSMNVFECLEC 577
QY 618 PVAGTAETQNTGBAGSKAC-QDQLSPTWSEIIEEDLRACFAGBAPLKDFSE 667
Db 578 ERANFPFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 613

RESULT 8

US-09-953-351-25

; Sequence 25, Application US/09953351

; Publication No. US20030036643A1

; GENERAL INFORMATION:

; APPLICANT: Li, Min

; APPLICANT: Melander, Christian

; APPLICANT: Liu, Hong-Xiang

; APPLICANT: Jin, Cheng He

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FUSION

; FILE REFERENCE: A-70814/RFT/RMS/RMK

; CURRENT APPLICATION NUMBER: US/09/953,351

; PRIOR FILING DATE: 2001-09-14

; PRIOR APPLICATION NUMBER: US 60/232,960

; PRIOR FILING DATE: 2000-09-14

; NUMBER OF SEQ ID NOS: 56

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 25

; LENGTH: 627

; TYPE: PRT

; ORGANISM: Barbarie duck parvovirus

US-09-953-351-25

Query Match 12.6%; Score 459; DB 10; Length 627;

Best Local Similarity 26.0%; Pred. No. 6.9e-31;

Matches 169; Conservative 78; Mismatches 264; Indels 140; Gaps 23;

QY 75 WEIT-----VDEMTKKQVFIFDSLKKCLFEVLNTKNIF-----PGDVNWVQHEWGK 122
Db 45 WEPTGIWNMEHVNLPMVTLADKI-----KNIFIQRMNQFNQDETDFFFLQEEGS 93

QY 123 DQGMHCHVLIGGKDFSOAQGKWRRLQNVYWSRWLTACNVQLTPAERIK---LREIAED 179
Db 94 EY-IHLHCCI-----AQNVRSFVLGRYMS-----QIKDSILRDVYEG 130

QY 180 N-----EWVTLTYKHQTKQYKCVLFGNMIAYYFLTKKISTSPRGGGYFLSSDSG 234
Db 131 KQVKIPDWFSTITKRGQGNKVTAA-----YILHYLIPKKQ-----PELQWAFNTMPLF 180

QY 235 WKTNFKEGEHLVSKLYTDDMRPEVTTVTTAQETKGRIGRIQTKKEVSIKTKLKELVHK 294
Db 181 TAAALCLQKQELLDAFQESEMNAVQEDQAAPL-----ISNRAAKNYSNLDVLIEM 236

QY 295 RVTSPEDEMMQPDYSYEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSKLT 354
Db 237 GITSEKQMLTENKESYRSFQATSSNNRQVKAALENABEMLLTATDYLGKDPVLDT 296

QY 355 NFSLPDTRTCRIFAFGHWNVYKCHAI CCVNLNRQGGKRNALFHPGPASTGKSIIAQIAQ 414
Db 297 K-----NRIYQILKLNYNPQYVGSVLCGWVKEFNKRNAILWYPATTKTNIABIAH 351

QY 415 AVNGVGCYNAANVFPNDCTNKNLIWEEAGNFGQOVNQFKCAICSGQTTIRIDOKGKSK 474
Db 352 AVPFYGCWNTNENFPFNDVCDKMLIWWBEGKMTNKVESAAILGGSVAVRDQCKGVS 411

RESULT 9

US-10-080-376-25

; Sequence 25, Application US/10080376

; Publication No. US20020172968A1

; GENERAL INFORMATION:

; APPLICANT: Li, Min

; APPLICANT: Bahiyat, Basail I.

; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES

; FILE REFERENCE: A-70295-2/RFT/RMS/RMK

; CURRENT APPLICATION NUMBER: US/10/080,376

; CURRENT FILING DATE: 2000-02-19

; PRIOR APPLICATION NUMBER: US 09/792,630

; PRIOR FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 87

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 25

; LENGTH: 627

; TYPE: PRT

; ORGANISM: Barbarie duck parvovirus

US-10-080-376-25

Query Match 12.6%; Score 459; DB 13; Length 627;

Best Local Similarity 26.0%; Pred. No. 6.9e-31;

Matches 169; Conservative 78; Mismatches 264; Indels 140; Gaps 23;

QY 75 WEIT-----VDEMTKKQVFIFDSLKKCLFEVLNTKNIF-----PGDVNWVQHEWGK 122
Db 45 WEPTGIWNMEHVNLPMVTLADKI-----KNIFIQRMNQFNQDETDFFFLQEEGS 93

QY 123 DQGMHCHVLIGGKDFSOAQGKWRRLQNVYWSRWLTACNVQLTPAERIK---LREIAED 179
Db 94 EY-IHLHCCI-----AQNVRSFVLGRYMS-----QIKDSILRDVYEG 130

QY 180 N-----EWVTLTYKHQTKQYKCVLFGNMIAYYFLTKKISTSPRGGGYFLSSDSG 234
Db 131 KQVKIPDWFSTITKRGQGNKVTAA-----YILHYLIPKKQ-----PELQWAFNTMPLF 180

QY 235 WKTNFKEGEHLVSKLYTDDMRPEVTTVTTAQETKGRIGRIQTKKEVSIKTKLKELVHK 294
Db 181 TAAALCLQKQELLDAFQESEMNAVQEDQAAPL-----ISNRAAKNYSNLDVLIEM 236

QY 295 RVTSPEDEMMQPDYSYEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSKLT 354
Db 237 GITSEKQMLTENKESYRSFQATSSNNRQVKAALENABEMLLTATDYLGKDPVLDT 296

QY 355 NFSLPDTRTCRIFAFGHWNVYKCHAI CCVNLNRQGGKRNALFHPGPASTGKSIIAQIAQ 414
Db 297 K-----NRIYQILKLNYNPQYVGSVLCGWVKEFNKRNAILWYPATTKTNIABIAH 351

QY 415 AVNGVGCYNAANVFPNDCTNKNLIWEEAGNFGQOVNQFKCAICSGQTTIRIDOKGKSK 474
Db 352 AVPFYGCWNTNENFPFNDVCDKMLIWWBEGKMTNKVESAAILGGSVAVRDQCKGVS 411

QY 475 QIEPTPVIMTNNITVVRIGCEERPEHTQPIIDRMLNIHLTHLPDGLVDKNE---- 530
Db 412 CIEPTVITNTDMCMIVDGNSTTMEHRIPLEERMFQIVLSHKLGEFGFKISKEVKEF 471

Db 412 CIEPTPVIIISNTDMCMIVDGNSTTMEHRIPLEERMFQIVLSHKLBNFGKISKEVKEF 471

Qy 531 --W-----PMICAWLVKNGYQSTMASYCAKWKGVDPWSEN-----WAEPK----- 568

Db 472 FKWANDNLVPVSEFKVPTNEQTKLTE-----PVPERANEPSPPKIWAPPTREBLEEI 525

Qy 569 -----VPTPINLGSARSPTTPKSTPLSONYALTPLASDLELALPEWSTPNT 617

Db 526 LRASPELFASVAPLP-----SSPDTSPKPKTKRGEYQVRCAMHSL-DNSMNVFECLEC 577

Qy 618 PVAGTAETONTGEAGSKAC-ODGQLSPTWSEIEDLRACFGAEPLKXDFSE 667

Db 578 ERANPEFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 613

RESULT 10

US-10-082-671-31

; Sequence 31, Application US/10082671

; Publication No. US20030049647A1

; GENERAL INFORMATION:

; APPLICANT: DAHIYAT, BASSIL

; APPLICANT: LI, MIN

; TITLE OF INVENTION: USE OF NUCLEIC ACID LIBRARIES TO CREATE TOXICOLOGICAL

; FILE OF INVENTION: PROFILES

; FILE REFERENCE: XEN/001

; CURRENT APPLICATION NUMBER: US/10/082,671

; CURRENT FILING DATE: 2002-05-17

; PRIOR APPLICATION NUMBER: 60/270,781

; PRIOR FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 58

; SOFTWARE: Patent in Ver. 2.1

; SEQ ID NO 31

; LENGTH: 627

; TYPE: PRT

; ORGANISM: Barbarie duck parvovirus

US-10-082-671-31

Query Match 12.6%; Score 459; DB 14; Length 627;
Best Local Similarity 26.0%; Pred. No. 6.9e-31;
Matches 169; Conservative 78; Mismatches 264; Indels 140; Gaps 23;

Qy 75 WETT-----VDENTKKQVFIQDSLVKKCLFEVLNTKNIF-----PGDVNMFVQHEWGK 122

Db 45 WEPTGIWNEHVNLPMTLADKI-----KNIFIQRWQNFQDDETFDFQLEEGS 93

Qy 123 DQGWCHVLIGGKDFSOAQGKWRRLQNVYWSRLVTACNVQLTPAERIK---LREIAD 179

Db 94 EY-IHLHCCI-----AQGNVRSFVLGRYMS-----QIKDSILRDVYEG 130

Qy 180 N-----EWVTLTYKHQTKDYTKCVLPFGNNIAYVFLTKKISTSPPRDGGYFLSSDSG 234

Db 131 KQVKIPDWFISITTKRGGQKNTVTA-----YILHVLIPKKQ-----PELQWAFNTMPLF 180

Qy 235 WKTNFLKEGERHLVSKLYTDDMRPETVETVTTAQTKEGRIOTKKEVSIKTLKELVHK 294

Db 181 TAAALCLQKROELLDAFOSEMNNAVQDOASTAAPL-----ISNRAAKNSVNLVDWLIEM 236

Qy 295 RVTSPEDEMMQPDSEIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSKLT 354

Db 237 GITSEKQWLTENKESYRSFQATSSNNRQVKALENARAEMLLTATDYLIGKDPVLDT 296

Qy 355 NFSLPDTRTCRIFAFHGWNVYVCHAI CCVLNRQGGKRNVLPHGPASTGKSTIAQAIQ 414

Db 297 K-----NRIYQILKLNNYNPQYVGVLCGWVKEEFNKNAINWLYGPATTKTNIAEIAH 351

Qy 415 AVGNVGCYNAANVPNDCTNKNLIWVEEAGNFGQOVNOFKAI CSQTIRIDOKGKSK 474

Db 352 AVFYGCYNNWTNENFPNDQVCKMLIWEEGKWTNKVVEESAKAILGGSAVRVDQCKGKV 411

Qy 475 QIEPTPVIMTTNENITVVRIGCEERPEHTOPIDRMLNIHLTHLPDGLVDKNE---- 530

Db 412 CIEPTPVIIISNTDMCMIVDGNSTTMEHRIPLEERMFQIVLSHKLBNFGKISKEVKEF 471

Qy 531 --W-----PMICAWLVKNGYQSTMASYCAKWKGVDPWSEN-----WAEPK----- 568

Db 472 FKWANDNLVPVSEFKVPTNEQTKLTE-----PVPERANEPSPPKIWAPPTREBLEEI 525

Qy 569 -----VPTPINLGSARSPTTPKSTPLSONYALTPLASDLELALPEWSTPNT 617

Db 526 LRASPELFASVAPLP-----SSPDTSPKPKTKRGEYQVRCAMHSL-DNSMNVFECLEC 577

Qy 618 PVAGTAETONTGEAGSKAC-ODGQLSPTWSEIEDLRACFGAEPLKXDFSE 667

Db 578 ERANPEFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 613

RESULT 11

US-10-097-100-25

; Sequence 25, Application US/10097100

; Publication No. US20030068649A1

; GENERAL INFORMATION:

; APPLICANT: Li, Min

; APPLICANT: Melander, Christian

; APPLICANT: Liu, Hong-Xiang

; APPLICANT: Jin, Cheng He

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FUSION LI

; FILE REFERENCE: A-70814/RFT/RMS/RMK

; CURRENT APPLICATION NUMBER: US/10/097,100

; CURRENT FILING DATE: 2002-03-12

; PRIOR APPLICATION NUMBER: US/09/953,351

; PRIOR FILING DATE: 2001-09-14

; PRIOR APPLICATION NUMBER: US 60/232,960

; PRIOR FILING DATE: 2000-09-14

; NUMBER OF SEQ ID NOS: 56

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 25

; LENGTH: 627

; TYPE: PRT

; ORGANISM: Barbarie duck parvovirus

US-10-097-100-25

Query Match 12.6%; Score 459; DB 14; Length 627;
Best Local Similarity 26.0%; Pred. No. 6.9e-31;
Matches 169; Conservative 78; Mismatches 264; Indels 140; Gaps 23;

Qy 75 WETT-----VDENTKKQVFIQDSLVKKCLFEVLNTKNIF-----PGDVNMFVQHEWGK 122

Db 45 WEPTGIWNEHVNLPMTLADKI-----KNIFIQRWQNFQDDETFDFQLEEGS 93

Qy 123 DQGWCHVLIGGKDFSOAQGKWRRLQNVYWSRLVTACNVQLTPAERIK---LREIAD 179

Db 94 EY-IHLHCCI-----AQGNVRSFVLGRYMS-----QIKDSILRDVYEG 130

Qy 180 N-----EWVTLTYKHQTKDYTKCVLPFGNNIAYVFLTKKISTSPPRDGGYFLSSDSG 234

Db 131 KQVKIPDWFISITTKRGGQKNTVTA-----YILHVLIPKKQ-----PELQWAFNTMPLF 180

Qy 235 WKTNFLKEGERHLVSKLYTDDMRPETVETVTTAQTKEGRIOTKKEVSIKTLKELVHK 294

Db 181 TAAALCLQKROELLDAFOSEMNNAVQDOASTAAPL-----ISNRAAKNSVNLVDWLIEM 236

Qy 295 RVTSPEDEMMQPDSEIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSKLT 354

Db 237 GITSEKQWLTENKESYRSFQATSSNNRQVKALENARAEMLLTATDYLIGKDPVLDT 296

Qy 355 NFSLPDTRTCRIFAFHGWNVYVCHAI CCVLNRQGGKRNVLPHGPASTGKSTIAQAIQ 414

Db 297 K-----NRIYQILKLNNYNPQYVGVLCGWVKEEFNKNAINWLYGPATTKTNIAEIAH 351

Qy 415 AVGNVGCYNAANVPNDCTNKNLIWVEEAGNFGQOVNOFKAI CSQTIRIDOKGKSK 474

Db 352 AVFYGCYNNWTNENFPNDQVCKMLIWEEGKWTNKVVEESAKAILGGSAVRVDQCKGKV 411

Qy 475 QIEPTPVIMTTNENITVVRIGCEERPEHTOPIDRMLNIHLTHLPDGLVDKNE---- 530

Db 412 CIEPTPVIIISNTDMCMIVDGNSTTMEHRIPLEERMFQIVLSHKLBNFGKISKEVKEF 471

QY 531 --W-----PMICAWLVKNGYQSTMASYCAKWKGVDPWSEN-----WAEPK----- 568
Db 472 FKWANDNLVPVSEFKVPTNEQTKLTE-----PVPERANEPSEPPKIWAPPTREELSEEI 525
QY 569 -----VPTPINLGSARSPTTPKSTPLSQNYALTPLASDLEDLALPWPSTPNT 617
Db 526 LRASPELFASVAPLP-----SSPDTSPKPKTRGEYQVRCAMHSL-DNSMNVFECLEC 577
QY 618 PVAGTATONTGAGSKAC-QDQOLSPTWSEIEEDLRACFGAPLKKDFSE 667
Db 578 ERANFPEFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 613

RESULT 12
US-10-023-208-25
; Sequence 25, Application US/10023208
; Publication No. US20030124537A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Daiyvat, Bassil I.
; TITLE OF INVENTION: PROCARYOTIC LIBRARIES AND USES
; FILE REFERENCE: A-70174-1/RT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/023,208
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/256,163
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Barbarie duck parvovirus
US-10-023-208-25

Query Match 12.6%; Score 459; DB 14; Length 627;
Best Local Similarity 26.0%; Pred. No. 6.9e-31;
Matches 169; Conservative 78; Mismatches 264; Indels 140; Gaps 23;

QY 75 WETT-----VDEMTKKQVIFDLSLVKKCLFEVLNTKNIF-----PGDVNWFVQHEWGK 122
Db 45 WEPTGIWMEHVLNPMWTLADKI-----KNIFQWQNFQODETDFPQLEEGS 93

QY 123 DQGWCHVLIGGKDFGQAQKQWRRQLNVYWSRWLTACNVQTPAERIK---LREIAD 179
Db 94 EY-IHLHCCI-----AQGNVRSFVLGRYMS-----QIKDSILRDVYEG 130

QY 180 N-----EWVLLTYKHQTKDYTKCVLFGNMAYYFLTKKISTSPRDGGYFLSDSG 234
Db 131 KQVKIPDMFSITTKRGQNKVTAA-----YILHLYLPKKQ-----PELQMAFTNPLF 180

QY 235 WKTNPLKEGGRHLVSKLYTDDMRPETVTTTAAQETKRGRIQTKKEVSIKTKLKVHK 294
Db 181 TAAALCLOKQELDLAFQESMAVQEDQASTAPL-----ISNRAKNVSNLVDMLIEM 236

QY 295 RVTSPEMMQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLT 354
Db 237 GITSEKQWLTENKESYRSFQATSSNNRQVKAALENARAEMLLTKTATDYLGKDPVLDT 296

QY 355 NFSLPDTRTCRIFAFHGWNTVYKCHAICCVLNROGGKRNNAVLFHGPASTGKSI 414
Db 297 K-----NRIYQILKLANNYPQYGVSLCGWVKEFNKRNAIWLYPATTKTNAIAEIAH 351

QY 415 AVGNVGCYNAANVFPNDCTNKNLIWVEAGNFQGVQVQFKAICSGQTIIRIDOKGKSK 474
Db 352 AVFPYGCNVNTNENFPNDCTNKNLIWVEAGNFQGVQVQFKAICSGQTIIRIDOKGKSK 411

QY 475 QIBPTPVIMTNNITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNE----- 530
Db 412 CIEPTPVIIISNTDMCMVDGNSTTMEHRIPLERMFQIVLSKLGNGFKISKKEVKEP 471

QY 531 --W-----PMICAWLVKNGYQSTMASYCAKWKGVDPWSEN-----WAEPK----- 568

Db 472 FKWANDNLVPVSEFKVPTNEQTKLTE-----PVPERANEPSEPPKIWAPPTREELSEEI 525
QY 569 -----VPTPINLGSARSPTTPKSTPLSQNYALTPLASDLEDLALPWPSTPNT 617
Db 526 LRASPELFASVAPLP-----SSPDTSPKPKTRGEYQVRCAMHSL-DNSMNVFECLEC 577
QY 618 PVAGTATONTGAGSKAC-QDQOLSPTWSEIEEDLRACFGAPLKKDFSE 667
Db 578 ERANFPEFQSLGE---NFCNQHG-----WYD-----CAFCNELKDDMNE 613

RESULT 13
US-09-792-630-21
; Sequence 21, Application US/09792630
; Patent No. US20020168640A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Daiyvat, Bassil I.
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES
; FILE REFERENCE: A-70295/RT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/792,630
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 610
; TYPE: PRT
; ORGANISM: adeno-associated virus 5
US-09-792-630-21

Query Match 12.5%; Score 453.5; DB 9; Length 610;
Best Local Similarity 27.2%; Pred. No. 2e-30;
Matches 157; Conservative 84; Mismatches 221; Indels 115; Gaps 22;

QY 67 WDQSEDMEWETTVDENTKKQVIFDLSLVKKCLFEVLNTKNIPFGDVNWFVQHEWGKDOGW 126
Db 35 WELPPSDNLTLVE---QPOLTVDARRIRRVFLTE---WNKFSQESKFFVQFKGSEY-F 88

QY 127 HCHVLIGGKDF-SQAQGWRRQLNVYWSRWLTACNVQTPAERIKLREI-AEDNEWVT 184
Db 89 HLTLVETSGISSMWLGRY-----VSQIRAOVLV---KVVFQIEQINDWA 132

QY 185 LLTYKHQTKDYTKCVLFGNMAYYFLTKKISTSPRDGGYFLSDSGSWKTNFLKEGE 244
Db 133 I-----TKVKGKANKVDSGYIAYLLPK-----VQPELOMAWTLNDE-- 171

QY 245 RHLVSKLYTDDMRPETVTTTAAQETKRGRIQTKKEVSIKTKLKE-----LV 292
Db 172 -YKLAALNLEERKRLVAQFLAESQSQSQAASQ--REFSADPVIKSKTSQKYMALVNLV 228

QY 293 HKRVTSPEMMQPDYSIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSK 352
Db 229 EHGITSEKQWLTENKESYLSFNSSTGNSRSQIKAAALONATKIMSLTKSADY----- 279

QY 353 LTNFSLPD---TRTCRIFAFHGWNTVYKCHAICCVLNROGGKRNNAVLFHGPASTGK 405
Db 280 LVGSSVPEDISKRNRIWQIFEMNGYDPAYAGSILYGC---QRSFNKRNVTWLYGPATTKG 336

QY 406 SITAQAIAGNVGCVNAANVFPNDCTNKNLIWVEAGNFQGVQVQFKAICSGQTIIR 465
Db 337 TNAIEAIAHTVPPYGCNVNTNENFPNDCTNKNLIWVEAGNFQGVQVQFKAICSGQTIIR 396

QY 466 IDQKGGSKQIEPTPVIMTNNITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGL 525
Db 397 VDQKCHSSVQIDSTPVIIVTSNTMVCVVDGNSTTTFHQQLPLEDMFKFELTKELPDPFGK 456

QY 526 VDKNEWPMTCAMLVKNGYQSTMASYCAKWKGVDPWSE-----NNAEPKVPTPI-- 573
Db 457 ITKQEVKDFFAW-----AKVQNVPTHEFKVPRELAKTGAESKLKRLPLGD 502

QY 574 -----NLLGSARSPT--TPKSTPLSQNYA-LTPL 600
Db 503 VTNTSYKSLKRLARLSFVPETPRSSDVTDPAPLRL 539

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:27:18 ; Search time 22.3333 Seconds
(without alignments)
2895.119 Million cell updates/sec

Title: US-10-069-056-14
Perfect score: 1637
Sequence: 1 MAGNAYSDEVLGATNWLKEK.....RACFGAEPLKKDFSEPLNLD 672

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3623	99.6	672	1 UYPVIM	noncapsid protein
2	3556	97.8	721	1 UYPVIM	noncapsid protein
3	3326	91.4	672	1 UYPV1	noncapsid protein
4	3322	91.3	668	1 A44276	noncapsid protein
5	2647	72.8	668	1 UYPVME	noncapsid protein
6	2644	72.7	668	1 UYPVFP	noncapsid protein
7	2642	72.6	668	1 UYPVCP	noncapsid protein
8	2485	68.3	660	1 UYPVPP	noncapsid protein
9	2476	68.1	662	1 UYPVNA	noncapsid protein
10	1621.5	44.6	392	1 UYPVIF	noncapsid protein
11	551	15.1	641	2 S41439	gene NS-1 protein
12	545.5	15.0	641	2 S41434	gene NS-1 protein
13	544	15.0	620	1 UYPVAP	noncapsid protein
14	534.5	14.7	641	2 S41861	gene NS-1 protein
15	459.5	12.6	626	2 S52209	noncapsid protein
16	419.5	11.5	536	1 UYADIA	noncapsid protein
17	350	9.6	726	1 UYPVSI	noncapsid protein
18	343	9.4	671	1 UYPV19	noncapsid protein
19	164	4.5	490	2 T44050	hypothetical prote
20	163.5	4.5	490	2 J01630	noncapsid protein
21	139	3.8	276	2 S26428	hypothetical 31.5K
22	133.5	3.7	849	1 UYPVAD	noncapsid protein
23	132	3.6	545	2 B44054	orf2 protein - Jun
24	124.5	3.4	647	1 W1L139	E1 protein - human
25	120.5	3.3	825	2 G96665	protein F22C12.12
26	116.5	3.2	614	1 W1L41	E1 protein - human
27	114.5	3.1	1098	2 E70232	hypothetical prote
28	113.5	3.1	631	2 S36505	E1 protein - human
29	113.5	3.1	647	2 S36517	E1 protein - human

ALIGNMENTS

RESULT 1

UYPVIM

noncapsid protein NS1 - minute virus of mice

C;Species: minute virus of mice, murine parvovirus

C;Date: 14-Nov-1983 #sequence_revision 28-Aug-1985 #text_change 09-Jul-2004

C;Accession: A03696

R;Astell, C.R.; Thomson, M.; Merchlinsky, M.; Ward, D.C.

Nucleic Acids Res. 11, 999-1018, 1983

A;Title: The complete DNA sequence of minute virus of mice, an autonomous parvovirus.

A;Reference number: A03696; MUID:83143341; PMID:6298737

A;Accession: A03696

A;Molecule type: DNA

A;Residues: 1-672 <AST>

A;Cross-references: UNIPROT:P03134; EMBL:V01115

C;Superfamily: parvovirus noncapsid protein

C;Keywords: noncapsid protein

Query Match 99.6%; Score 3623; DB 1; Length 672;

Best Local Similarity 99.7%; Pred. No. 2.2e-255;

Matches 670; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	1	MAGNAYSDEVLGATNWLKEKSNQEVFSFVKENVQLNGKDIGNSYKKSLQDELKSLQ	60
Db	1	MAGNAYSDEVLGATNWLKEKSNQEVFSFVKENVQLNGKDIGNSYKKSLQDELKSLQ	60
Qy	61	RGAEITWDQSEDMEWETTVDEMTKKQVFI	120
Db	61	RGAEITWDQSEDMEWETTVDEMTKKQVFI	120
Qy	121	GKQDQGHCHVLI	180
Db	121	GKQDQGHCHVLI	180
Qy	181	EWTLTYKHKQKDYKCVLFGNMIAYFLKXKIS	240
Db	181	EWTLTYKHKQKDYKCVLFGNMIAYFLKXKIS	240
Qy	241	KEGERHLVSKLYTDDMRP	300
Db	241	KEGERHLVSKLYTDDMRP	300
Qy	301	DWMWQDPSYIEMMAQPGGENLKNL	360
Db	301	DWMWQDPSYIEMMAQPGGENLKNL	360
Qy	361	TRTCRIFAFHGMNVKVC	420
Db	361	TRTCRIFAFHGMNVKVC	420
Qy	421	CYNAANVFPNDCTNKLWVEAGNFGQOVNQK	480
Db	421	CYNAANVFPNDCTNKLWVEAGNFGQOVNQK	480

probable RNA helic
hypothetical prote
E1 protein - human
probable ATP bindi
structural polypro
transcription regu
type I site-specif
dyatrophin - mouse
E1 protein - bovin
E1 protein - human
hypothetical WW do
tipC protein - ali
hypothetical prote
hypothetical prote
E1 protein - human
large T antigen -


```
Qy 421 CYNAAVNFPNDCTNKNLIWVEAGNFGQOVNQFKAI CSQTIRIDQKSGSKQIEPTP 480
|||||
Db 421 CYNAAVNFPNDCTNKNLIWVEAGNFGQOVNQFKAI CSQTIRIDQKSGSKQIEPTP 480
|||||
Qy 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
|||||
Db 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
|||||
Qy 541 NGYQSTWASCAKWKGVDPDSENWABPKVTPPINLLGSARSPTTPKSTPLSQNYALTPL 600
|||||
Db 541 NGYQSTWASCAKWKGVDPDSENWABPKVTPPINLLGSARSPTTPKSTPLSQNYALTPL 600
|||||
Qy 601 ASLEDLALEPWSPTNPVAGTAETONTGAGSKACODGQGLSPTWSIEIEDLRACFAEP 660
|||||
Db 601 ASLEDLALEPWSPTNPVAGTAETONTGAGSKACODGQGLSPTWSIEIEDLRACFAEP 660
|||||
Qy 661 LKXDFSEPLNLD 672
|||: |||
Db 661 LESDFNEELTLD 672
|||: |||

RESULT 4
A44276
noncapsid protein NS1 - parvovirus LuIII
C:Species: parvovirus LuIII
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C:Accession: A44276
R:Difford, N.; Chen, K.C.; Bates, R.C.; Lederma, M.
Virology 192, 339-345, 1993
A:Title: The complete nucleotide sequence of parvovirus LuIII and localization of a unique
A:Reference number: A44276; MUID:93297126; PMID:8517025
A:Accession: A44276
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-668 <DIF>
A:Cross-references: UNIPROT:P36311; GB:M81898
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match 91.3%; Score 3322; DB 1; Length 668;
Best Local Similarity 91.4%; Pred. No. 1.7e-233;
Matches 614; Conservative 23; Mismatches 31; Indels 4; Gaps 2;

Qy 1 MAGNAYSDEVLGATNWLKEKSNQEVFSFVKENNVQNGDKIGNYSYKKELODELKSLQ 60
|||||
Db 1 MAGNAYSDEVLGATNWLKDKSNQEVFSFVKENNVQNGDKIGNYSYKKELODELKSLQ 60
|||||
Qy 61 RGAETTDQSEDMEWETTVDEMTHKKQVFI FDSL VKKCLFEVLNTKNI FPGDVNMFVQHEW 120
|||||
Db 61 RGAETTDQSEDMEWESSVDELTKQVFI FDSL VKKCLFEVLNTKNI FPGDVNMFVQHEW 120
|||||
Qy 121 GKQGWCHVLI GKGDFSOAGKWRRLQNVYWSRWLVLTACNVQLTTPAERIKLREIAEDN 180
|||||
Db 121 GKQGWCHVLI GKGDFSOAGKWRRLQNVYWSRWLVLTACNVQLTTPAERIKLREIAEDQ 180
|||||
Qy 181 EWTLLTYKHQTKKDYTKCVFNGMIAYFLTKKKISTSPRDGGYFLSSDSGWKTNFL 240
|||||
Db 181 EWTLLTYKHQTKKDYTKCVFNGMIAYFLTKKKICTSPRDGGYFLSSDSGWKTNFL 240
|||||
Qy 241 KEGERHLVSKLYTDDMRPEVETVTTAQTGKGRIOTKKEVSIKTTLKL VHKRVTSPE 300
|||||
Db 241 KEGERHLVSKLYTDDMRPEVETVTTAQTGKGRIOTKKEVSIKTTLKL VHKRVTSPE 300
|||||
Qy 301 DNMWMPDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSKLTNFSLPD 360
|||||
Db 301 DNMWMPDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSKLTNFSLPD 360
|||||
Qy 361 TRTCRIFAFHGWNYKVCHAI CCVNLNROGGKRNVLPHGPASTGKSI IAAQIAQAVGNV 420
|||||
Db 361 TRTCRIFAFHGWNYKVCHAI CCVNLNROGGKRNVLPHGPASTGKSI IAAQIAQAVGNV 420
|||||
Qy 421 CYNAAVNFPNDCTNKNLIWVEAGNFGQOVNQFKAI CSQTIRIDQKSGSKQIEPTP 480
|||||
```

```
Db 421 CYNAAVNFPNDCTNKNLIWVEAGNFGQOVNQFKAI CSQTIRIDQKSGSKQIEPTP 480
|||||
Qy 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
|||||
Db 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
|||||
Qy 541 NGYQSTWASCAKWKGVDPDSENWABPKVTPPINLLGSARSPTTPKSTPLSQNYALTPL 600
|||||
Db 541 NGYQSTWASCAKWKGVDPDSENWABPKVTPPINLLGSARSPTTPKSTPLSQNYALTPL 600
|||||
Qy 601 ASLEDLALEPWSPTNPVAGTAETONTGAGSKACODGQGLSPTWSIEIEDLRACFAEP 660
|||||
Db 598 -SLEDLALEPWSPTNPVAGTAETONTGAGSKACODGQGLSPTWSIEIEDLRACFAEP 660
|||||
Qy 661 LKXDFSEPLNLD 672
|||: |||
Db 657 WKSQSEQLPNLD 668
|||: |||

RESULT 5
UYPVME
noncapsid protein NS1 - mink enteritis virus (strain Abashiri)
C:Species: mink enteritis virus, MEV
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: A38350
R:Kariatsumari, T.; Horiuchi, M.; Hama, E.; Yaguchi, K.; Ishiguro, N.; Goto, H.; Shinag
J. Gen. Virol. 72, 867-875, 1991
A:Title: Construction and nucleotide sequence analysis of an infectious DNA clone of the
A:Reference number: A38350; MUID:91202123; PMID:2016597
A:Accession: A38350
A:Molecule type: DNA
A:Residues: 1-668 <XAR>
A:Cross-references: UNIPROT:P27438; GB:D00765; NID:g222435; PIDN:BAA00662.1; PID:g222436
C:Superfamily: parvovirus noncapsid protein
C:Keywords: noncapsid protein

Query Match 72.8%; Score 2647; DB 1; Length 668;
Best Local Similarity 73.1%; Pred. No. 2.1e-184;
Matches 491; Conservative 71; Mismatches 104; Indels 6; Gaps 5;

Qy 1 MAGNAYSDEVLGATNWLKEKSNQEVFSFVKENNVQNGDKIGNYSYKKELODELKSLQ 60
|||||
Db 1 MSGNQYTEEWEGVNNLKKHAENAFSFKCNVQNGDKDVHNNYTKPIQNEELTSLI 60
|||||
Qy 61 RGAETTDQSD--BDMEMETTVDEMTHKKQVFI FDSL VKKCLFEVLNTKNI FPGDVNMFVQH 118
|||||
Db 61 RGAQTAMDQTEEBEEMDSEVDSLAKKQVQTPDALLKKCLFEVFSKNI EPNECVWFIOH 120
|||||
Qy 119 EWGKDGWCHVLI GKGDFSOAGKWRRLQNVYWSRWLVLTACNVQLTTPAERIKLREIAE 178
|||||
Db 121 EWGKDGWCHVLI LHSKNLQQAATGKWLRRQNMNYSRWLVLTCLSVNLTPTEKILREIAE 180
|||||
Qy 179 DNEWVTLTYKHQTKKDYTKCVLFGNMIAYFLTKKKISTSPRDGGYFLSSDSGWKTN 238
|||||
Db 181 DSEWVTLTYRHQTKKDYVYVMVHFGNMIAYFLTKKKI-VHMTKESGYFLSDSGWKFN 239
|||||
Qy 239 FLKEGERHLVSKLYTDDMRPEVETVTTAQTGKGRIOTKKEVSIKTTLKL VHKRVTS 298
|||||
Db 240 FMKYQDRHTVSTLYTEQMKPEVETVTTAQTGKGRIOTKKEVSIKTTLRLVSKRVTS 299
|||||
Qy 299 PEDWMMQPDYSYIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKAETSKLTNFSL 358
|||||
Db 300 PEDWMMQPDYSYIEMMAQPGGENLLKNTLEICTLTARTKTAFLDILEKANNTKLTNFDL 359
|||||
Qy 359 PDTRTCRIFAFHGWNYKVCHAI CCVNLNROGGKRNVLPHGPASTGKSI IAAQIAQAVGN 418
|||||
Db 360 ANSKTCQIFMHHGWNVIKVCHAIACVNLNROGGKRNVLPHGPASTGKSI IAAQIAQAVGN 419
|||||
Qy 419 VGCYNAANVNFPPNDCTNKNLIWVEAGNFGQOVNQFKAI CSQTIRIDQKSGSKQIEP 478
|||||
Db 420 VGCYNAANVNFPPNDCTNKNLIWVEAGNFGQOVNQFKAI CSQTIRIDQKSGSKQIEP 479
|||||
Qy 479 TPVIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWL 538
|||||
```

Db 480 TPVIMTNNITIVRIGCEERPEHTQPIRDRMLNKLVLCKLPGLVDRKEEPLICAWL 539
Qy 539 VKNGYQSTWASYCAKWKGVDPWSNWAEPKVPPTPINLLGSARSPTTTPKSTPLSQNVALT 598
Db 540 VKHGYSSTWANYTHHWKVPWEDNWAEPKIQEGI-CKOLETQAASNPQSDHVL 598
Qy 599 PLASDLLEALPWSPTNTPVAGTAETQNTGEAGSKACQDQGLSPTWSEIEEDLRACFGA 658
Db 599 PLTPDVVVDLALPWSPTDTPIAETA-NQOSNQLG-VTHKDVQASPTWSEIEADLRIFTS 656
Qy 659 EPLKDFSEPLN 670
Db 657 EQLEEDFRDLD 668

RESULT 6

UYVFP
noncapsid protein NS1 - feline panleukopenia virus (strain 193)
C;Species: feline panleukopenia virus, FPLV
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004
C;Accession: A36608
R;Martyn, J.C.; Davidson, B.E.; Studdert, M.J.
J. Gen. Virol. 71, 2747-2753, 1990
A;Title: Nucleotide sequence of feline panleukopenia virus: comparison with canine parvovirus
A;Reference number: A36608; MUID:91073139; PMID:2174965
A;Accession: A36608
A;Molecule type: DNA
A;Residues: 1-668 <MAR>
A;Cross-references: UNIPROT:P24842; GB:X55115; NID:G60863; PIDN:CAA38910.1; PID:G60864
C;Superfamily: parvovirus noncapsid protein
C;Keywords: noncapsid protein

Query Match 72.7%; Score 2644; DB 1; Length 668;
Best Local Similarity 73.1%; Pred. No. 3.4e-184;
Matches 491; Conservative 71; Mismatches 104; Indels 6; Gaps 5;
Qy 1 MAGNAYSDEVLGATNWLKESNQEVFVFKENNVOLNGKDIGNWSYKKELODELKSLQ 60
Db 1 MSGNQYTEEVGVNWLKHAENEAFVFKCDNVOLNGKDVWRNNYTKPIQNEELTSLI 60
Qy 61 RGAETWDS--EDMEWETTVDENTKKQVIFDLSLVKKCLFEVLNTKNIIPGDVNVFVQH 118
Db 61 RGAQTAMDQTEEEEDWSEVDSLAKKQVQTFDALIKKCLFEVFSKNIEPNECVWFIQH 120
Qy 119 EWGKDGQWCHVLIGGKDFSAQGWRRQLNVYWSRWLVTAQNVLTTPAERIKLREIAE 178
Db 121 EWGKDGQWCHVLHSHKLNQQAATGKWLRRQNMWYWSRWLVTLCSVNLTPTEKIKLREIAE 180
Qy 179 DNEWVTLTYKHQTKKDYTKCVLFGNMIAIYFLTKKKISTSPRDGGYFLSSDSGKTN 238
Db 181 DSEWVTILTYRHQTKKDYKVMVHFGNMIAIYFLTKKKI-VHMTKESGYFLSDSGWKN 239
Qy 239 FLKEGERHLVSKLYTDDMRPETVETVTTTAQETKRGRIQTKKEVSIKTLKELVHKRVTS 298
Db 240 FMKYQDRQIVSTLYTEQMKPEVETVTTTAQETKRGRIQTKKEVSIKTLRDLVSKRVTS 299
Qy 299 PEDMMQPDQSYTEMMAQPGENLLKNTLEICTLTARTKTAPELILEKADNTKLTNFDL 358
Db 300 PEDMMQPDQSYTEMMAQPGENLLKNTLEICTLTARTKTAPELILEKADNTKLTNFDL 359
Qy 359 PDTRTCRIFAFHGMNYYKVCCHAICCVLNROGGKRNVLPHGPASTGKSIQAIAQAVGN 418
Db 360 ANSRCTQIFRMHGMNWKVCHAIACVLNROGGKRNVLPHGPASTGKSIQAIAQAVGN 419
Qy 419 VGCYNAANVFPNDCTNKNLIWVEEAGNFQGVNQFKAICSGQTIIRIDQKGGSKQIEP 478
Db 420 VGCYNAANVFPNDCTNKNLIWVEEAGNFQGVNQFKAICSGQTIIRIDQKGGSKQIEP 479
Qy 479 TPVIMTNNITIVRIGCEERPEHTQPIRDRMLNHLTHLPDGLVDRKNEPMLICAWL 538
Db 480 TPVIMTNNITIVRIGCEERPEHTQPIRDRMLNKLVLCKLPGLVDRKEEPLICAWL 539

Qy 539 VKNGYQSTWASYCAKWKGVDPWSNWAEPKVPPTPINLLGSARSPTTTPKSTPLSQNVALT 598
Db 540 VKHGYSSTWANYTHHWKVPWEDNWAEPKIQEGI-CKOLETQAASNPQSDHVL 598
Qy 599 PLASDLLEALPWSPTNTPVAGTAETQNTGEAGSKACQDQGLSPTWSEIEEDLRACFGA 658
Db 599 PLTPDVVVDLALPWSPTDTPIAETA-NQOSNQLG-VTHKDVQASPTWSEIEADLRIFTS 656
Qy 659 EPLKDFSEPLN 670
Db 657 EQLEEDFRDLD 668

RESULT 7

UYVCP
noncapsid protein NS1 - canine parvovirus (strain N)
C;Species: canine parvovirus, CPV
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
C;Accession: A29962
R;Reed, A.P.; Jones, E.V.; Miller, T.J.
J. Virol. 62, 266-276, 1988
A;Title: Nucleotide sequence and genome organization of canine parvovirus.
A;Reference number: A29962; MUID:88062992; PMID:2824850
A;Accession: A29962
A;Molecule type: DNA
A;Residues: 1-668 <REE>
A;Cross-references: UNIPROT:P12929; EMBL:M19296; NID:G333438; PIDN:AAA67459.1; PID:G33343
C;Superfamily: parvovirus noncapsid protein
C;Keywords: noncapsid protein

Query Match 72.6%; Score 2642; DB 1; Length 668;
Best Local Similarity 73.1%; Pred. No. 4.8e-184;
Matches 491; Conservative 71; Mismatches 104; Indels 6; Gaps 5;
Qy 1 MAGNAYSDEVLGATNWLKESNQEVFVFKENNVOLNGKDIGNWSYKKELODELKSLQ 60
Db 1 MSGNQYTEEVGVNWLKHAENEAFVFKCDNVOLNGKDVWRNNYTKPIQNEELTSLI 60
Qy 61 RGAETWDS--EDMEWETTVDENTKKQVIFDLSLVKKCLFEVLNTKNIIPGDVNVFVQH 118
Db 61 RGAQTAMDQTEEEEDWSEVDSLAKKQVQTFDALIKKCLFEVFSKNIEPNECVWFIQH 120
Qy 119 EWGKDGQWCHVLIGGKDFSAQGWRRQLNVYWSRWLVTAQNVLTTPAERIKLREIAE 178
Db 121 EWGKDGQWCHVLHSHKLNQQAATGKWLRRQNMWYWSRWLVTLCSVNLTPTEKIKLREIAE 180
Qy 179 DNEWVTLTYKHQTKKDYTKCVLFGNMIAIYFLTKKKISTSPRDGGYFLSSDSGKTN 238
Db 181 DSEWVTILTYRHQTKKDYKVMVHFGNMIAIYFLTKKKI-VHMTKESGYFLSDSGWKN 239
Qy 239 FLKEGERHLVSKLYTDDMRPETVETVTTTAQETKRGRIQTKKEVSIKTLKELVHKRVTS 298
Db 240 FMKYQDRQIVSTLYTEQMKPEVETVTTTAQETKRGRIQTKKEVSIKTLRDLVSKRVTS 299
Qy 299 PEDMMQPDQSYTEMMAQPGENLLKNTLEICTLTARTKTAPELILEKADNTKLTNFDL 358
Db 300 PEDMMQPDQSYTEMMAQPGENLLKNTLEICTLTARTKTAPELILEKADNTKLTNFDL 359
Qy 359 PDTRTCRIFAFHGMNYYKVCCHAICCVLNROGGKRNVLPHGPASTGKSIQAIAQAVGN 418
Db 360 ANSRCTQIFRMHGMNWKVCHAIACVLNROGGKRNVLPHGPASTGKSIQAIAQAVGN 419
Qy 419 VGCYNAANVFPNDCTNKNLIWVEEAGNFQGVNQFKAICSGQTIIRIDQKGGSKQIEP 478
Db 420 VGCYNAANVFPNDCTNKNLIWVEEAGNFQGVNQFKAICSGQTIIRIDQKGGSKQIEP 479
Qy 479 TPVIMTNNITIVRIGCEERPEHTQPIRDRMLNHLTHLPDGLVDRKNEPMLICAWL 538
Db 480 TPVIMTNNITIVRIGCEERPEHTQPIRDRMLNKLVLCKLPGLVDRKEEPLICAWL 539
Qy 539 VKNGYQSTWASYCAKWKGVDPWSNWAEPKVPPTPINLLGSARSPTTTPKSTPLSQNVALT 598
Db 540 VKHGYSSTWANYTHHWKVPWEDNWAEPKIQEGINSPG-CKOLETQAASNPQSDHVL 598

Db 421 YNAANVFPNDCTNKLWIEEAGNSQVNOQKAIKCSGTIRIDQKSGSKQIEPTPV 480
Qy 482 IMTNENITVVICERPERPHTQPIRDRMLNIHLTHLPGDGLVDKNWPMICAWLVKN 541
Db 481 IMTNEDITKVICERPERPHTQPIRDRMLNINUTRKLPGDGLLEETWEPFLICAWLVKK 540
Qy 542 GYOSTMASYCAKMGKVPDWSENWAEKVPPTPINLLGSARSPFTTPKSTPLSONYALTPLA 601
Db 541 GYQATMASYHHWGNVDPWSEKWEKVPQPTINTPTDSQIS-TSVKTSPADNNYAAATPIQ 599
Qy 602 SDLE-DLALPSPWTPNTPVAGTAQTONTGAGSKACODGQL---SPTWSEIEDLRACFG 657
Db 600 EDLDLALPSPWSEPTTPTTNLHLTPT-----PDSAIRTPSPWSEIEDLRACFG 652

RESULT 10
Query Match 44.6%; Score 1621.5; DB 1; Length 392;
Best Local Similarity 77.5%; Pred. No. 3.6e-110; Indels 3; Gaps 3;
Matches 306; Conservative 33; Mismatches 53; Indels 3; Gaps 3;
C;Species: feline panleukopenia virus (fragment)
C;Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 09-Jul-2004
C;Accession: A03697
R;Carlson, J.; Rushlow, K.; Maxwell, I.; Maxwell, F.; Winston, S.; Hahn, W.
J. Virol. 55, 574-587, 1985
A;Title: Cloning and sequence of DNA encoding structural proteins of the autonomous parvovirus
A;Reference number: A03697; MUID:85265017; PMID:2991581
A;Accession: A03697
A;Molecule type: DNA
A;Residues: 1-392 <CAR>
A;Cross-references: UNIPROT:P06431; EMBL:M10824; NID:G333474; PIDN:AAA47160.1; PID:G3334
C;Superfamily: parvovirus noncapsid protein
C;Keywords: noncapsid protein

Query Match 44.6%; Score 1621.5; DB 1; Length 392;
Best Local Similarity 77.5%; Pred. No. 3.6e-110; Indels 3; Gaps 3;
Matches 306; Conservative 33; Mismatches 53; Indels 3; Gaps 3;
C;Species: feline panleukopenia virus, FPLV
C;Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 09-Jul-2004
C;Accession: A03697
R;Carlson, J.; Rushlow, K.; Maxwell, I.; Maxwell, F.; Winston, S.; Hahn, W.
J. Virol. 55, 574-587, 1985
A;Title: Cloning and sequence of DNA encoding structural proteins of the autonomous parvovirus
A;Reference number: A03697; MUID:85265017; PMID:2991581
A;Accession: A03697
A;Molecule type: DNA
A;Residues: 1-392 <CAR>
A;Cross-references: UNIPROT:P06431; EMBL:M10824; NID:G333474; PIDN:AAA47160.1; PID:G3334
C;Superfamily: parvovirus noncapsid protein
C;Keywords: noncapsid protein

Qy 276 IQTKKESVSIKTLKELVHKVTSPEWMMQPSYIEMMAQPGGENLLKNTLEICTLTLA 335
Db 1 IQTKKESVSIKTLRLDLSKRVTSPEWMMQPSYIEMMAQPGGENLLKNTLEICTLTLA 60
Qy 336 RTKTAFLILEKARTSKLTNFSLPDTRCTIFAFHGNVYKVAICCVLNROGKKNV 395
Db 61 RTKTAFLILEKANNTKLTNFDLANSRTQIFRMHGNWIKVCHAIACVILNROGKKNV 120
Qy 396 LFHGPASTGKSIQAIAQAVGNGVCYNAANVPFNDCTNKLWIEEAGNFQGVNQF 455
Db 121 LFHGPASTGKSIQAIAQAVGNGVCYNAANVPFNDCTNKLWIEEAGNFQGVNQF 180
Qy 456 KAICSGQTIRIDQKSGSKQIEPTPVTMTNENITVVRIGCEERPEHTQPIRDRMLNIHL 515
Db 181 KAICSGQTIRIDQKSGSKQIEPTPVTMTNENITVVRIGCEERPEHTQPIRDRMLNIHL 240
Qy 516 THLPGDGLVDKNWPMICAWLVKNGYOSTMASYCAKMGKVPDWSENWAEKVPPTPINL 575
Db 241 VCKLPGDGLVDKEWPLICAWLVKHYESTMANYTHHWGPEDENWAEKPIQEGVNS 300
Qy 576 LGSARSPFTTPKSTPLSONYALTPLASDLEALPSPWTPNTPVAGTAQTONTGAGSKA 635
Db 301 PG-CKDLETOASNPQSDHVLPTPLDVLALPSPWTPNTPVAGTAQTONTGAGSKA 357
Qy 636 CQDQLSPTWSEIEDLRACFGAEPLKDFSEPLN 670
Db 358 HKDVQASPTWSEIEDLRACFGAEPLKDFSEPLN 670

RESULT 11
S41439
Gene NS-1 protein - Aleutian mink disease virus
C;Species: Aleutian mink disease virus
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: S41439
R;Gottschalk, E.; Alexandersen, S.; Storgaard, T.; Bloom, M.E.; Aasted, B.

submitted to the EMBL Data Library, January 1994
A;Description: Sequence comparison of the non-structural genes of four different types of
A;Reference number: S41434
A;Accession: S41439
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-641 <GOT>
A;Cross-references: UNIPROT:Q65017; EMBL:Z29576
C;Genetics:
A;Introns: 586/1
C;Superfamily: parvovirus noncapsid protein

Query Match 15.1%; Score 551; DB 2; Length 641;
Best Local Similarity 25.8%; Pred. No. 5.4e-32;
Matches 178; Conservative 108; Mismatches 262; Indels 142; Gaps 25;
Qy 18 KEKSNQEVFSFVKENNVQLNGKDIGWNSYKKELODE-----LKSQORGAETTD 68
Db 8 EQRRLOQLYVOLKKEIN---DGGVAVLFOQKYTYTDKNKPTKATPLRTTSSDLRLAFD 64
Qy 69 QSDMEWETTVDKMTK-----KOVFIQFDSLVKCLFEVLNTKNIFFPGDVNMFVQ 117
Db 65 SIEE-NLTASNEQLTNGINFCKLTGLKTLTLLLDKHVSHRWD--NNK-----VNLIWQ 115
Qy 118 HEWKGQOGHCHVLIG-----GKDFSOQAGKWRRLN-----VWWSRLVTACNVQ- 164
Db 116 IEKGTQOQFHIHCCLGYDFRNEDEPKDVQKSLG-WFMKRLNKLALYNNH-----HCDIOD 170
Qy 165 -LTPASRIKLRETAEDNEWVTLTYKHQTKDYTKVCLFNGNMIAYVFLTKKISPPR 223
Db 171 IKDPEKAKMLKVVVEDGPTPKYKFNQTKQYKPNKPVHLRDYTFIYLFNKKINDSM- 229
Qy 224 DGGYFLSSDSGKWTNFKEGEHLVSKLYTDDMRPETVTV-----TTAQ 269
Db 230 -DCYFAAGNGIVDN-LTNKERKTLKMYLDEQSSDIMDANIDWEDQDAPKVTDOTDSA 287
Qy 270 ETYRG-----RIQTKKEV-----SIKTLKELVHKRVTSPEWMMQ 307
Db 288 TTKTGTSLIWKSCATKVTSKKEVANFPVQPSKKLYSAQNTLDALFNVCFTPEDMIKOS 347
Qy 308 DSVIENMAOPGGENLKNLTLEICTLARTKTAFLDILE--KAETSKLTNFSLPDTRTCR 365
Db 348 KYLELSLENGPQKINTLLHNNQVKTSTMTAFDCCIENEEDDKPLLATIKDM---- 403
Qy 366 IFAFHGNVYKVAICCVLNROGKKNVLFHGPASTGKSIQAIAQAVGNGVCYNA 425
Db 404 -----GLNEQYLKVLCTILTQGGKRGCIWFYGPQGTGKTLASLICKATVNGVMTS 458
Qy 426 NVNFPNDCTNKLWIEEAGNFQGVNQFKAICSGQTIRIDQKSGSKQIEPTPMT 485
Db 459 NPNFPMTDCGNRIIWAECGNLGNVEDFKAITGGGVKVDTKNQPSIKGC-VIVTS 517
Qy 486 NENITVVRIGCEERPEHTQPIRDRMLNIHL-----THLPGDGLVDKNWPMICAWL 538
Db 518 NTNITKVTGCVETNAHAEPKMKIRCMKTINPKITPG-----MLKRWL 566
Qy 539 VKNGYOSTMASYCAKMGKVP-DWSENWAEKVPPTPINLLGSARSPFTTPKSTPLSONYAL 597
Db 567 -----NTWDRQPIQLSHEMPELYLET-----TGNSSATTATKNTGNSQ---- 605
Qy 598 TPLASDLEALPSPWTPNTPVAGTAQTQN 627
Db 606 PTTAKGAESVNTENCDDTPKRGASSVPPKQH 635

RESULT 12
S41434
Gene NS-1 protein - Aleutian mink disease virus
C;Species: Aleutian mink disease virus
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: S41434
R;Gottschalk, E.; Alexandersen, S.; Storgaard, T.; Bloom, M.E.; Aasted, B.
submitted to the EMBL Data Library, January 1994

A;Description: Sequence comparison of the non-structural genes of four different types of parvovirus
A;Reference number: S41434
A;Accession: S41434
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-641 <GOT>
A;Cross-references: UNIPROT:Q65020; EMBL:Z29577
C;Genetics:
A;Introns: 586/1
C;Superfamily: parvovirus noncapsid protein

Query Match 15.08; Score 545.5; DB 2; Length 641;
Best Local Similarity 25.28; Pred. No. 1.3e-31;
Matches 182; Conservative 110; Mismatches 260; Indels 171; Gaps 26;

Qy 18 KEKSNQEVFVFKNENVLQNGDKIGWNSYKKEIQEDELKSLQGAETWTD 68
Db 8 EQRRLQLFPE-KPKTE--VADGEGLAMLFQKQTYTDKDKPKTKATPLPRTTSDLRLAFD 64
Qy 69 QSEDMETVDEMTKKQVIFDLSLVKCKLFEVLNTKNI-----FPGDVNMFVQHEWKGKD 123
Db 65 SIEE-TLKTSNQCLTNNDINFCKLTGLKTL--VLLDKHVRSHRDANKVNFVQVEKGKT 121
Qy 124 QGWHCHVLIG-----GKDFSOAQGWRRQLN-----VYWSRWLVTAACVOLTPAERIK 172
Db 122 QQFHHICCLGYFDKDESDKDVQKSLG-WFKKLNKDLAVIWSNHHCDIOGIQSDSGRADN 180
Qy 173 LREIAEDNEWTLTYKHKOTKDYTKCVLFGNMIAYFLTKKISTSPRDG--GYFLS 230
Db 181 LKTIWEDGP-TKPKYFNKQTKQDYNKPTNLRDYLILYFNKOKIT-----QEGMDGYAS 235
Qy 231 SDGKWTNFKKEGRHVLVSKLYTDDM-----RPETVETTTVAQETPKRG 274
Db 236 GNGGIIDN-LTNKERRALKRWYLDQSQGILDEDWEDSQSAPKVDQDTSATSKTGS 294
Qy 275 -----RQTKKEVSI-----KTLTKELVHKRVTSPEDEMMQDPSYLEMM 314
Db 295 LVKSCATKVTSKKEVAIPVKQPSKQWTSQAQNTLDDLYMFCFTPEDMILKMSDYLEMS 354
Qy 315 AQPGENLLKNTLEICTLTARTKTAFDLILE--KAETSKLTNFSLPDTRTCRIFAPHGW 372
Db 355 LEPNGAQKINTLLHMQVRTSSNLTAPECIKFNEEDEKPLDITKDM-----GL 405
Qy 373 NYKVKCHAI CCVLRQGGKRNALVFGHPASTGKSIIAQIAQAVGNVGCYNAANVPFPN 432
Db 406 NEQHLKVLCTILTQSGKRGCVWFYFGGTGKTLASLICKAVNYGMVTTSNPNPFWT 465
Qy 433 DCTNKNLIWVEEAGNFQGVNQVQKAI CSQTIRIDQKGSKOIEPTFVIMTNENITVV 492
Db 466 DCGNRNIWAEECGNIGNYVEDEKAITGGGVKVDTKNKPQSIKGS-VIVTSNTNITKV 524
Qy 493 RIGCEERPEHTPIRDRMLNIHL-----THLPGDFGLVDKNWPMICAWLVGNGYQS 545
Db 525 TVGCVETNVAEPLKQRMVKIRCMKPNPKTKITPG-----MLKTIW----- 566
Qy 546 TMASYCAKWKGVDPWSENWAEKVPVTPINLLGSARSPTTPKSTPLSQNYALTPLASDLE 605
Db 567 -----STWDRIIP-----IKLSHEMP 581
Qy 606 DLALPEWSTNTPVAGTAETQNTCEAGSKACQDQGLSPTWSEIHEEDLRACFGEPLKQF 665
Db 582 ELYLET-SGPNSS-SATTATKST-----GNLQPTTAETAESVNTANCDDTPKRGAS 628
Qy 666 SEP 668
Db 629 SVP 631

RESULT 13
UYVPAP
noncapsid protein NS1 - Aleutian mink disease virus (strain ADV-G)
N;Alternate names: left-ORF protein
C;Species: Aleutian mink disease virus

C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C;Accession: A36760; A35529
R;Bloom, M.E.; Alexandersen, S.; Perryman, S.; Lechner, D.; Wolfmberger, J.B.
J. Virol. 62, 2903-2915, 1988
A;Title: Nucleotide sequence and genomic organization of Aleutian mink disease parvovirus
A;Reference number: A36760; MUID:88275062; PMID:2839709
A;Accession: A36760
A;Molecule type: DNA
A;Residues: 1-620 <BL2>
A;Cross-references: UNIPROT:P24030; EMBL:M20036
C;Superfamily: parvovirus noncapsid protein
C;Keywords: noncapsid protein

Query Match 15.08; Score 544; DB 1; Length 620;
Best Local Similarity 26.68; Pred. No. 1.7e-31;
Matches 157; Conservative 104; Mismatches 229; Indels 100; Gaps 18;

Qy 18 KEKSNQEVFVFKNENVLQNGDKIGWNSYKKEIQEDELKSLQGAETWTDQSE-DNEWE 76
Db 38 EQRRLQLVQLKKEIN---DGEGLVAMLFQKQTYTDKDKPKTKATPLPRTTSSDLRLAFD 94
Qy 77 TTVDENMTKKQVIFDLSLVKCKLFEVLNTKNI-----FPGDVNMFVQHEWKGQDQGW 126
Db 95 SIEENLTASNEHLTNNEINFCKLTGLKTLILLIDKHVKSHEWDSKNVNLIIQIEKGTQOF 154
Qy 127 HCHVLIG-----GKDFSOAQGWRRQLN-----VYWSRWLVTAACVQ--LTPAERIKL 173
Db 155 HIHCLGLGYFDKDESDKDVQKSLG-WFKMRLNKLAVIYSN---HHCDIQIDKDPEDRAKN 210
Qy 174 RETAEDNEWTLTYKHKOTKDYTKCVLFGNMIAYFLTKKISTSPRDGCVFLSSDS 233
Db 211 LKVIWIEDGPTPKYKFNKQTKQDYNKPVHLRDYTFIYFNKDKINTDSM--DGYFAAGNG 268
Qy 234 GWTNFKKEGRHVLVSKLYTDDMRPETVETV-----TTAQTETKRG----- 274
Db 269 GIVDN-LTNKERTKRWYLDQSDIMDANIDWEDQDAPKVDQDTSATTKGTSLIW 327
Qy 275 -----RQTKKEV-----SIKTTLKELVHKRVTSPEDEMMQDPSYIEMMAQP 317
Db 328 KSCATKVTSKKEVANPVQPSKSLYSAQSLDALFNVCFTPEDMIKQSDKYLELSLEP 387
Qy 318 GGENLLKNTLEICTLTARTKTAFDLILE--KAETSKLTNFSLPDTRTCRIFAPHGWYV 375
Db 388 NGPQKINTLLHMQVKTSTMTAPDCIIFKNEBEDDKPLLATIKDM-----GLNEQ 438
Qy 376 KVCHAI CCVLRQGGKRNALVFGHPASTGKSIIAQIAQAVGNVGCYNAANVPFPNDCT 435
Db 439 YLKKVLCITLTQSGKRGCVWFYFGGTGKTLASLICKATVNYGMVTTSNPNPFWTDCG 498
Qy 436 NKNLIIWVEEAGNFQGVNQVQKAI CSQTIRIDQKGSKOIEPTFVIMTNENITVVRI 495
Db 499 NRNIWAEECGNIGNYVEDEKAITGGGVKVDTKNKPQSIKGC-VIVTSNTNITKVTVG 557
Qy 496 CEEPRPEHTPIRDRMLNIHL-----THLPGDFGLVDKNWPMICAWL 538
Db 558 CVETNAHAELPKQRMVKIRCMKTNPKTKITPG-----MLKRWL 596

RESULT 14
S41861
gene NS-1 protein - Aleutian mink disease virus
C;Species: Aleutian mink disease virus
C;Date: 15-Sep-1994 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
C;Accession: S41861; S41436
R;Gottschalk, E.; Alexandersen, S.; Storgaard, T.; Bloom, M.E.; Aasted, B.
submitted to the EMBL Data Library, January 1994
A;Description: Sequence comparison of the non-structural genes of four different types of
A;Reference number: S41434
A;Accession: S41861
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-641 <GOT>
A;Cross-references: UNIPROT:Q65023; EMBL:X77085; NID:9452597; PIDN:CAA54355.1; PID:945252

C;Genetics:

A;Introns: 586/1

C;Superfamily: parvovirus noncapsid protein

Query Match 14.7%; Score 534.5; DB 2; Length 641;
Best Local Similarity 25.7%; Pred. No. 8.5e-31;
Matches 178; Conservative 106; Mismatches 250; Indels 159; Gaps 27;

QY 18 KEKSNQEVFSVFVFNENVLNGKIDGWSYKKELEQDE-----LKSQRGAETWTD 68
DB 8 EOKRLQDLYTQKKE---VADGEGLAWLFOQKTYTDKDKNPKTKATPRLTSSDLRLAFD 64
QY 69 QSEDM-----EWETTVEMT-----KKQVFIFDSLVKCKLFEVLNTKNIPFGDVNWFVQ 117
DB 65 SIEETLIASNEW-LTKDBEINFCKLTGLGKTLVLVDKHKVSHRWA-----DKINFIQ 115
QY 118 HEWKQDQWHCHVLIG-----GKFSQAQGWRRQLNVYWSRWLTACNVOLTPAER 170
DB 116 IEKGTQHFHHCCLGYFDKNEPKDVQKSLG-WLIKKLN-----RDLAALFNSHHCDQD 170
QY 171 IKLREIAEDN-EWV-----TLLTYKHQTKDYTKCVLFGNMIAYYFLTKKISTSPPR 223
DB 171 IKDPEAKANLKVIEDGPTPKYKYNKQTKQDYNKPTLSDYVILYFNKDKIH-----K 226
QY 224 DG--GYFLSSDSGWKTNFLKEGERHLVSKLYTDDMRPETVE-----TTVTT 267
DB 227 EGMGYVAAAGNGLIDN-LTNKERKALRKMYLDQSSDIMDADIDWEDGQDAPKVTQTD 285
QY 268 AQETKRG-----RIQTKKEV-----SIKTLKELVHKVTSPEWMM 305
DB 286 SATSKTGTSLIWKSCATKVTSKKEVANPVQOPSKLYSAQNTLDALFNVGCFPTEDMIK 345
QY 306 QPDSYIEMMAOPGGENLLKNTLEICTLTARTKTAFLILE--KAETSCLTNFSLPDR 363
DB 346 QSDXYLELSELPNGPKINTILLHNQVKTSTMTAFDCLIKFNEEEDKFLATIKDM-- 403
QY 364 CRIFAFHGNVYKCHAI CCVLNROGGRNAVLFPHGPASTGKSIIAQIAQAVNGVCYN 423
DB 404 -----GLNEQYLLKVLCTILTQGGKRGCIWFGPGGTGLLASLICKATVNYGMVT 456
QY 424 AANYNFPNDCTNKLWVEAGNFGQOVNQFKAI CSQTIRIDOKGSKQIEPTPVM 483
DB 457 TSNENFPWTDCGRNIIWABECGNLGNWVEDFKAITGGGVKVDTKNKPQSKIGC-VIV 515
QY 484 TTENITVVRIGCEERPEHTQPIRDRMLNIHL-----THLPGDFGLVDKNEWPMICA 536
DB 516 TSNITITVTVGCVETWAHAEPKQRMKIRCMKTINPKTI TPG-----MLKR 564
QY 537 WLKXNGYQSTWASYCAKWKV-----DWSENWAEKVPKPTPINLLGSARSPP-TTPKSTP 590
DB 565 WL-----NTWDRQPIQLSHEMPELYLETTGPN-----SSATSATKTTGNSQP 606
QY 591 LSONYALTPLASDLELALPWPSTPNTPVAGTA 623
DB 607 TTAATBSVSTADCD-----TPKRGAS 628

RESULT 15

S52209

noncapsid protein NS1 - muscovy duck parvovirus

N;Alternate names: nonstructural protein NS1; NS1 protein

C;Species: muscovy duck parvovirus

C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004

C;Accession: S52209

R;Zadori, Z.; Erdei, J.; Nagy, J.; Kisari, J.

submitted to the EMBL Data Library, September 1993

A;Reference number: S52209

A;Accession: S52209

A;Molecule type: DNA

A;Residues: 1-626 <ZAD>

A;Cross-references: UNIPROT:Q83288; EMBL:X75093; PIDN:CAA52983.1; PID:g6090

A;Experimental source: strain FM

C;Genetics:

A;Gene: NS1

C;Superfamily: parvovirus noncapsid protein

C;Keywords: noncapsid protein

Query Match 12.6%; Score 459.5; DB 2; Length 626;
Best Local Similarity 25.3%; Pred. No. 2.3e-25;
Matches 165; Conservative 82; Mismatches 263; Indels 141; Gaps 22;

QY 75 WETT-----VDEMTKKOVIFDSLVKCKLFEVLNTKNIF-----PGDVNWFVQHEWGK 122
DB 45 WETFGIWNMEHVNPLWPTLADKI-----KNIFQWRNQFNQDETDFFFLQEEGS 93
QY 123 DQGMCHCVLIGKDFSOAQGWRRQLNVYWSRWLTACNVOLTPAERIK---LREIAED 179
DB 94 EY-IHLHAVCPGECRPFVLGRY-----MSQIKDSILRDVYEG 129
QY 180 N-----EWTLLTYKHQTKDYTKCVLFGNMIAYYFLTKKISTSPDRDGGVFLSDSG 234
DB 130 KQVKIPDMFWSITTKRGGQNKVTAA-----YTLHYLIPKKQ-----PELQWAFNMPLF 179
QY 235 WKTNFKEGEGERHLVSKLYTDDMRPETVEITVTTAQTCKRGRIQTKKEVSIKTTLKEIVHK 294
DB 180 TAAALCLQKQKQELDAFOSEMNNAVQEDQASTAPL-----ISNRAKNYSNVLVDLIEM 235
QY 295 RVTSPEDDMMWQPDYSYIEMMAOPGGENLLKNTLEICTLTARTKTAFLILEKAETSKLT 354
DB 236 GIITSEKQWLTENKESYRSFOATSSNNRQVKAALENARAEMLLTKTATDYLIGKDPVLDT 295
QY 355 NPSLPDTRICRIPAFHGNVYKCHAI CCVLNROGGRNAVLFPHGPASTGKSIIAQIAQ 414
DB 296 K-----NRIYQILKLNYPQVYVGVLCGWVKREFNKRNAIMLYGPATTKTNIAEAIH 350
QY 415 AVGNVCYNAANVFPNDCTNKLWVEAGNFGQOVNQFKAI CSQTIRIDOKGSK 474
DB 351 AVPFYCVVNTNENFPNDCTNKLWVEAGNFGQOVNQFKAI CSQTIRIDOKGSKGV 410
QY 475 QIEPTFVIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNE---- 530
DB 411 CIEPTFVIITNTDMCMIVDGNSTTMEHRI PLEERMFQIVLSHKLGNFGKISKKEVKEF 470
QY 531 --W-----PMICAWLVKNGYQSTWASYCAKWKVDPDWSN-----WAEPK----- 568
DB 471 FKWANDNLVPVVSSEFKVPTNEQTKLTE-----PVPERANEPESEPPKI WAPPTRELEEI 524
QY 569 -----VPTPINLLGSARSPTTPKSTPLSQNYALTPLASDLELALPWPSTPNT 617
DB 525 LRASPELPASVAPLP-----SSPDTSPARKKTRGEYQVRCAMHSL-DNSMNVFCLC 576
QY 618 PVAGTAETQNTGBAGSKAC-QDQQLSPWTWSEIIEEDLRACFGAPPLKKDFSE 667
DB 577 ERANPFEPQSLGE---NFCNQHG-----WYD-----CAFCNELKDDVNE 612

Search completed: January 22, 2005, 03:42:03

Job time : 24.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 03:19:28 ; Search time 106.333 Seconds
(without alignments)
3636.223 Million cell updates/sec

Title: US-10-069-056-14

Perfect score: 3637

Sequence: 1 MAGNAYSDEVLTGNWLKEK.....RACFGAEPLKKDFSEPLNLD 672

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3623	99.6	672	1 VNC5 MUMIV	P03134 murine minu
2	3623	99.6	721	2 Q84365	Q84365 murine minu
3	3554	97.7	721	1 VNC5 MUMIM	P07300 murine minu
4	3554	97.7	721	2 Q84363	Q84363 murine minu
5	3552	97.7	672	2 Q83429	Q83429 mouse parvo
6	3345	92.0	672	2 Q8JVL4	Q8JVL4 rat minute
7	3336	91.7	672	2 Q8JVL8	Q8JVL8 hamster rat
8	3335	91.7	672	2 P88899	P88899 kilham rat
9	3327	91.5	672	2 Q8JVL8	Q8JVL8 rat minute
10	3326	91.4	672	1 VNC5 PAVHH	P03133 hamster par
11	3323	91.4	672	2 Q8JVL6	Q8JVL6 rat minute
12	3322	91.3	668	1 VNC5 PAVL3	P36311 parvovirus
13	3303	90.8	665	2 Q71159	Q71159 kilham rat
14	2935.5	80.7	671	2 Q71157	Q71157 rat parvovi
15	2655	73.0	668	2 P89513	P89513 feline panl
16	2654	73.0	668	2 P89516	P89516 feline panl
17	2653	72.9	668	2 P89512	P89512 feline panl
18	2653	72.9	668	2 P89515	P89515 feline panl
19	2653	72.9	668	2 P90449	P90449 feline panl
20	2653	72.9	668	2 P90484	P90484 feline panl
21	2652	72.9	668	2 P89514	P89514 feline panl
22	2652	72.9	668	2 P90472	P90472 feline panl
23	2647	72.8	668	1 VNC5 MEVA	P27438 mink enteri
24	2644	72.7	668	1 VNC5 FPV19	P24842 feline panl
25	2642	72.6	668	1 VNC5 PAVCN	P12929 canine parv
26	2637	72.5	668	2 Q84393	Q84393 canine parv
27	2614	71.9	668	2 Q70M74	Q70M74 canine parv
28	2614	71.9	668	2 CAE47433	CAE47433 canine pa
29	2485	68.3	660	1 VNC5 PAVPN	P18547 porcine par
30	2476	68.1	662	1 VNC5 PAVPK	P52502 porcine par
31	2472	68.0	662	2 Q6RED5	Q6RED5 porcine par

32	2472	68.0	662	2 Q6TPD8	Q6TPD8 porcine par
33	2472	68.0	662	2 AAQ90279	AAQ90279 porcine p
34	2472	68.0	662	2 AAR91039	AAR91039 porcine p
35	2463	67.7	662	2 Q6PS60	Q6PS60 porcine par
36	2463	67.7	662	2 AAS93262	AAS93262 porcine p
37	1933	53.1	397	2 Q993M6	Q993M6 autonomous
38	1621.5	44.6	392	1 VNC5 FPV	P06431 feline panl
39	551	15.1	641	2 Q65017	Q65017 aleutian mi
40	545.5	15.0	641	2 Q65020	Q65020 aleutian mi
41	544	15.0	590	1 VNC5 ADVG	P24030 aleutian mi
42	543	14.9	620	2 Q96607	Q96607 aleutian mi
43	534.5	14.7	641	2 Q65023	Q65023 aleutian mi
44	460.5	12.7	610	2 Q6JL80	Q6JL80 bovine aden
45	460.5	12.7	610	2 AAR26464	AAR26464 bovine ad

ALIGNMENTS

```
RESULT 1
VNC5 MUMIV
ID VNC5 MUMIV STANDARD; PRT; 672 AA.
AC P03134;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVPL1).
GN Name=NS1;
OS Murine minute virus (Murine parvovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10794;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83143341; PubMed=6298737;
RA Astell C.R.; Thomson M.; Merchlinsky M.; Ward D.C.;
RT "The complete DNA sequence of minute virus of mice, an autonomous
parvovirus."
RL Nucleic Acids Res. 11:999-1018(1983).
RN [2]
RP HOMOLOGIZATION, AND INTERACTION WITH SYNCRIP.
RX MEDLINE=99102562; PubMed=9847309;
RA Harris C.E.; Boden R.A.; Astell C.R.;
RT "A novel heterogeneous nuclear ribonucleoprotein-like protein
interacts with NS1 of the minute virus of mice."
RL J. Virol. 73:72-80(1999).
CC -I- FUNCTION: Seems necessary for viral DNA replication.
CC -I- SUBUNIT: Forms homooligomers. Interacts with human SYNCRIP.
CC -I- DOMAIN: The N-terminus (residues 1-275) possesses a negative effect
on transactivation.
CC -I- DOMAIN: The C-terminus (residues 543-672) possesses an activation
domain.
CC -I- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL; J02275; AAA67109.1; -.
DR EMBL; V01115; AAA24309.1; ALT_INIT.
DR PIR; A03696; UYFVIM.
DR TRANSFAC; T02375; -.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; 1.
DR ATP-binding; DNA replication; Noncapsid protein;
KW Nonstructural protein.
KW DOMAIN 1 276 Interacts with SYNCRIP.
FT NP_BIND 399 406 ATP (Potential).
SQ SEQUENCE 672 AA; 76248 MW; 50298F27662E3C1D CRC64;
```

Query Match	99.6%; Score 3623; DB 1; Length 672;
Best Local Similarity	99.7%; Pred. No. 3.5e-259;
Matches 670; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
Qy 1	MAGNAYSDEVLGATNWLKESNQEVFSFVKENVQLNGKDIGWNSYKKELOEBELSKLQ 60
Db 1	MAGNAYSDEVLGATNWLKESNQEVFSFVKENVQLNGKDIGWNSYKKELOEBELSKLQ 60
Qy 61	RGAEATTWQSEDMEWETTVDENTKKQVIFDLSLVKKCLFEVLNTKNIIPPGDVNVFQHEW 120
Db 61	RGAEATTWQSEDMEWETTVDENTKKQVIFDLSLVKKCLFEVLNTKNIIPPGDVNVFQHEW 120
Qy 121	GKDQGHCHVLITGGKDFSOAQGWRRQLNVYWSRWLVACNVQLTPAERIKLEIAEDN 180
Db 121	GKDQGHCHVLITGGKDFSOAQGWRRQLNVYWSRWLVACNVQLTPAERIKLEIAEDN 180
Qy 181	EWTLTYKHQTKDYTKCVLFGNMIAFYFLTKKISTSPRDGGYFLSSDSGWKTNFL 240
Db 181	EWTLTYKHQTKDYTKCVLFGNMIAFYFLTKKISTSPRDGGYFLSSDSGWKTNFL 240
Qy 241	KEGERHLVSKLYTDDMRPETVETVTTAQETKRGRIQTKKEYSIKITLKVHVRVTSPE 300
Db 241	KEGERHLVSKLYTDDMRPETVETVTTAQETKRGRIQTKKEYSIKITLKVHVRVTSPE 300
Qy 301	DWMMPQDSYIENMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLNFSLPD 360
Db 301	DWMMPQDSYIENMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLNFSLPD 360
Qy 361	TRTCRIFAHGWNVYKCHAI CCVLNROGKKNVLFHGPASTGKSIITAOAIAQAVGNV 420
Db 361	TRTCRIFAHGWNVYKCHAI CCVLNROGKKNVLFHGPASTGKSIITAOAIAQAVGNV 420
Qy 421	CYNAANVNFNDCTNKNLIWVEEAGNFQOQVNOQFKAICSGOTIRIDQKGSKQIEPTP 480
Db 421	CYNAANVNFNDCTNKNLIWVEEAGNFQOQVNOQFKAICSGOTIRIDQKGSKQIEPTP 480
Qy 481	VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAMLVK 540
Db 481	VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAMLVK 540
Qy 541	NGYQSTMASYCAKWKVPDWSNWAEPKVPPTINLLGSARSPTTPKSTPLSQNYALTPL 600
Db 541	NGYQSTMASYCAKWKVPDWSNWAEPKVPPTINLLGSARSPTTPKSTPLSQNYALTPL 600
Qy 601	ASLDELALPEWSTPNTPVAGTAETONTGEAGSKACQDQGLSPTWSEIEEDLRACFGAEP 660
Db 601	ASLDELALPEWSTPNTPVAGTAETONTGEAGSKACQDQGLSPTWSEIEEDLRACFGAEP 660
Qy 661	LKKDFSEPLNLD 672
Db 661	LKKDFSEPLNLD 672
RESULT 2	
Q84365	PRELIMINARY; PRT; 721 AA.
ID	Q84365
AC	Q84365
DT	01-NOV-1996 (TrEMBLrel. 01, Created)
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE	Nonstructural protein.
GN	Name=NS1.
OS	Murine minute virus (Murine parvovirus).
OC	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX	NCBI_TaxID=10794;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=MVM;
RC	MEDLINE=83143341; PubMed=6298737;
RA	Astell C.R., Thomson M., Merchlinsky M., Ward D.C.;
RA	"The complete DNA sequence of minute virus of mice, an autonomous
RT	parvovirus.";
RL	Nucleic Acids Res. 11:999-1018(1983).

```
Qy 661 LKQDFSEPLNLD 672
Db 710 LKQDFSEPLNLD 721

RESULT 3
VNC5 MUMIM STANDARD; PRT; 672 AA.
AC P07300; P10837;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVP1).
GN Name=NS1;
OS Murine minute virus (strain MVM1) (Murine parvovirus).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10795;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86115415; PubMed=3502703;
RA Astell C.R., Gardiner E.M., Tattersall P.;
RT "DNA sequence of the lymphotropic variant of minute virus of mice,
RT MVM(1), and comparison with the DNA sequence of the fibrotropic
RT prototype strain.";
RL J. Virol. 57:656-669(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85242059; PubMed=3855242;
RA Sahli R., McMaster G.K., Hirt B.;
RT "DNA sequence comparison between two tissue-specific variants of the
RT autonomous parvovirus, minute virus of mice.";
RL Nucleic Acids Res. 13:3617-3633(1985).
CC -1- FUNCTION: Seems necessary for viral DNA replication.
CC -1- SUBUNIT: Forms homooligomers. Interacts with human SYNCRIP (By
CC similarity).
CC -1- DOMAIN: The N-terminus (residues 1-275) possesses a negative effect
CC on transactivation (By similarity).
CC -1- DOMAIN: The C-terminus (residues 543-672) possesses an activation
CC domain (By similarity).
CC -1- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X02481; -; NOT ANNOTATED_CDS.
CC DR EMBL; M12032; AAA69567.1; -.
CC DR InterPro; IPR001257; Parvo NS1.
CC DR Pfam; PF01057; Parvo NS1; 1.
CC KW ATP-binding; DNA replication; Noncapsid protein;
CC KW Nonstructural protein.
CC FT DOMAIN 1 276 Interacts with SYNCRIP (By similarity).
CC FT NP BIND 399 406 ATP (Potential).
CC FT CONFLICT 597 597 I -> L (in Ref. 2).
CC SQ SEQUENCE 672 AA; 76140 MW; 25F025FE32B84DF0 CRC64;

Query Match 97.7%; Score 3554; DB 1; Length 672;
Best Local Similarity 97.5%; Pred. No. 4.4e-254;
Matches 655; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

Qy 1 MAGNAYSDEVLGATNWLKESNVEFVFPKNIENVQLNGDKIGNYSYKELQEDLKSQ 60
Db 1 MAGNAYSDEVLGATNWLKESNVEFVFPKNIENVQLNGDKIGNYSYKELQEDLKSQ 60.
Qy 61 RGAETTTDQSEDMEWETTVDEMTKQVFIYDLSLVKCLFEVLATKNIFPGDVNWFVQHEW 120
Db 61 RGAETTTDQSEDMEWETTVDEMTKQVFIYDLSLVKCLFEVLATKNIFPGDVNWFVQHEW 120
Qy 121 GKQDQGHCHVLIIGKQFSDQAGKWRRLVYNSRWLVATCNVQLTTPAERIKLREIAEDN 180
```

```
QY 1 MAGNAYSDEVLGATNMLKEKSNQEVFSFVKFNENVLQNGKDIGNWSYKKELODELKSLQ 60
Db 50 MAGNAYSDEVLTGNTLMLKEKSNQEVFSFVKETEDVQNGKDIGNWSYKKELODELKSLQ 109
QY 61 RGAETTWDSQSEDMWETTTVDKMTKKQVFIYDLSLVKKCLFEVLNTKNIIFPGDVNMFVQHEW 120
Db 110 RGAETTWDSQSEDMWETTTVDKMTKKQVFIYDLSLVKKCLFEVLNTKNIAPADVTWVFQHEW 169
QY 121 GKQGHCHVLIGGKDFSOAQGWRRQLNYSRWLVACNVQLTPAERIKLREIAEDN 180
Db 170 GKQGHCHVLIGGKDFSOAQGWRRQLNYSRWLVACNVQLTPAERIKLREIAEDS 229
QY 181 EWTLLTYKHQKQDYTKCVLFQGNMIAYYFLTKKISTSPPRDGGYFLSSDGSWKTNFL 240
Db 230 EWTLLTYKHQKQDYTKCVLFQGNMIAYYFLTKKISTSPPRDGGYFLSSDGSWKTNFL 289
QY 241 KEGERHLVSKLYTDDMRPETVETTTAQTETKGRIOQTKKEVSIKTTLKELVHKRVTSPE 300
Db 290 KEGERHLVSKLYTDDMRPETVETTTAQTETKGRIOQTKKEVSIKTTLKELVHKRVTSPE 349
QY 301 DNMWMPDSYIEMMAOPGGENLLKNTLEICTLTARTKTAFDLILEKAETSCLTNFSLPD 360
Db 350 DNMWMPDSYIEMMAOPGGENLLKNTLEICTLTARTKTAFDLILEKAETSCLTNFSLPD 409
QY 361 TRTCRIFAFHGMNYSVKVCHACCVLNROGGKRNVLPHGPASTGKSIIAQAIQAAGVNGV 420
Db 410 TRTCRIFAFHGMNYSVKVCHACCVLNROGGKRNVLPHGPASTGKSIIAQAIQAAGVNGV 469
QY 421 CYNAAVNPFNDCTNKNLIWVEEAGNFGQVQNFKAICSGQTIRIDQKGGSKQIEPTP 480
Db 470 CYNAAVNPFNDCTNKNLIWVEEAGNFGQVQNFKAICSGQTIRIDQKGGSKQIEPTP 529
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNEWPMICAWLVK 540
Db 530 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNEWPMICAWLVK 589
QY 541 NGYQSTMASYCAKWKVPDSENWAEKVPPTPINLLGSARSPTTTPKSTPLSQNYALTPL 600
Db 590 NGYQSTMASYCAKWKVPDSENWAEKVPPTPINLLGSARSPTTTPKSTPLSQNYALTPL 649
QY 601 ASDLEDLALPEWSTPNTPVAGTAETONTGEGSKACODGQSLSTWSEIEDLRACFGAEP 660
Db 650 ASDLEDLALPEWSTPNTPVAGTAETONTGEGSKACODGQSLSTWSEIEDLRACFGAEP 709
QY 661 LKXDFSEPLNLD 672
Db 710 LKXDFSEPLNLD 721

RESULT 5
Q83429 ID Q83429 PRELIMINARY; PRT; 672 AA.
AC Q83429
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Nonstructural protein 1.
GN Name=NSI;
OS Mouse parvovirus 1.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=35340;
RN [1]
RP MEDLINE=94365951; PubMed=8083985;
RX Ball-Goodrich L.J., Johnson E.;
RA "Molecular characterization of a newly recognized mouse parvovirus.";
RL J. Virol. 68:6476-6486(1994).
DR EMBL; U12469; AAA61405.1;
DR GO; GO:0019012; C:virion; IEA.
DR GO; GO:0019079; P:virion genome replication; IEA.
DR InterPro; IPR001257; Parvo_NSI.
DR Pfam; PF01057; Parvo_NSI; I.
KW Nonstructural protein.
```

```
SQ SEQUENCE 672 AA; 76112 MW; 31C6365276727363 CRC64;
Query Match 97.7%; Score 3552; DB 2; Length 672;
Best Local Similarity 97.3%; Pred. No. 6.2e-254;
Matches 654; Conservative 10; Mismatches 8; Indels 0; Gaps 0;
QY 1 MAGNAYSDEVLGATNMLKEKSNQEVFSFVKFNENVLQNGKDIGNWSYKKELODELKSLQ 60
Db 1 MAGNAYSDEVLTGNTLMLKEKSNQEVFSFVKETEDVQNGKDIGNWSYKKELODELKSLQ 60
QY 61 RGAETTWDSQSEDMWETTTVDKMTKKQVFIYDLSLVKKCLFEVLNTKNIIFPGDVNMFVQHEW 120
Db 61 RGAETTWDSQSEDMWETTTVDKMTKKQVFIYDLSLVKKCLFEVLNTKNIAPADVTWVFQHEW 120
QY 121 GKQGHCHVLIGGKDFSOAQGWRRQLNYSRWLVACNVQLTPAERIKLREIAEDN 180
Db 121 GKQGHCHVLIGGKDFSOAQGWRRQLNYSRWLVACNVQLTPAERIKLREIAEDS 180
QY 181 EWTLLTYKHQKQDYTKCVLFQGNMIAYYFLTKKISTSPPRDGGYFLSSDGSWKTNFL 240
Db 181 EWTLLTYKHQKQDYTKCVLFQGNMIAYYFLTKKISTSPPRDGGYFLSSDGSWKTNFL 240
QY 241 KEGERHLVSKLYTDDMRPETVETTTAQTETKGRIOQTKKEVSIKTTLKELVHKRVTSPE 300
Db 241 KEGERHLVSKLYTDDMRPETVETTTAQTETKGRIOQTKKEVSIKTTLKELVHKRVTSPE 300
QY 301 DNMWMPDSYIEMMAOPGGENLLKNTLEICTLTARTKTAFDLILEKAETSCLTNFSLPD 360
Db 301 DNMWMPDSYIEMMAOPGGENLLKNTLEICTLTARTKTAFDLILEKAETSCLTNFSLPD 360
QY 361 TRTCRIFAFHGMNYSVKVCHACCVLNROGGKRNVLPHGPASTGKSIIAQAIQAAGVNGV 420
Db 361 TRTCRIFAFHGMNYSVKVCHACCVLNROGGKRNVLPHGPASTGKSIIAQAIQAAGVNGV 420
QY 421 CYNAAVNPFNDCTNKNLIWVEEAGNFGQVQNFKAICSGQTIRIDQKGGSKQIEPTP 480
Db 421 CYNAAVNPFNDCTNKNLIWVEEAGNFGQVQNFKAICSGQTIRIDQKGGSKQIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNEWPMICAWLVK 540
Db 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNHLTHLPGDFGLVDKNEWPMICAWLVK 540
QY 541 NGYQSTMASYCAKWKVPDSENWAEKVPPTPINLLGSARSPTTTPKSTPLSQNYALTPL 600
Db 541 NGYQSTMASYCAKWKVPDSENWAEKVPPTPINLLGSARSPTTTPKSTPLSQNYALTPL 600
QY 601 ASDLEDLALPEWSTPNTPVAGTAETONTGEGSKACODGQSLSTWSEIEDLRACFGAEP 660
Db 601 ASDLEDLALPEWSTPNTPVAGTAETONTGEGSKACODGQSLSTWSEIEDLRACFGAEP 660
QY 661 LKXDFSEPLNLD 672
Db 661 LKXDFSEPLNLD 672

RESULT 6
Q8JUV14 ID Q8JUV14 PRELIMINARY; PRT; 672 AA.
AC Q8JUV14;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Nonstructural protein 1.
GN Name=NSI;
OS Rat minute virus 1c.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=172387;
RN [1]
RP MEDLINE=22120170; PubMed=12124471;
RX Wan C.H., Soderlund-Venermo M., Pintel D.J., Riley L.K.;
RA "Molecular characterization of three newly recognized rat
RT parvoviruses.";
```



```
RL J. Gen. Virol. 83:2075-2083 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Wan C.-H., Soderlund-Venemo M., Pintel D., Riley L.K.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF332884; AM93279.1; -.
DR GO; GO:0019012; C:virion; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR001257; Parvo_NSL.
DR Pfam; PF01057; Parvo_NSL; 1.
KW Nonstructural protein.
SQ SEQUENCE 672 AA; 75988 MW; 52DF6549349CF3FD CRC64;

Query Match 92.0%; Score 3345; DB 2; Length 672;
Best Local Similarity 91.5%; Pred. No. 1.3e-238;
Matches 616; Conservative 22; Mismatches 34; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVLGATNMLKEKSNQEVFSFVKNNVQLNGKDIGNWSYKKELODELKSLQ 60
DB 1 MAGNAYSDEVLGATNMLKDKSSQEVFSFVKNNVQLNGKDIGNWSYKKELODELKSLQ 60
QY 61 RGAETTWQSDMEDWETTVDEMTHKQVFI FDSL VKKCLFEVLNTKNI FPGDVNMFVQHEW 120
DB 61 RGAETTWQSDMEDWESA VDDMTKQVFI FDSL VKKCLFEVLNTKNI APSDVTWVQHEW 120
QY 121 GKQGGHCHVLIGGKDFSOAQGWRRQLNLYWSRWLVTCNVQLTPAERIKLREIAEDN 180
DB 121 GKQGGHCHVLIGGKDFSOAQGWRRQLNLYWSRWLVTCNVQLTPAERIKLREIAEDS 180
QY 181 EWTLLTYKHKOTKDYTKCVLFQGNMIAYFLTKKISTSPPRDGGYFLSSDSGWKTNFL 240
DB 181 EWTLLTYKHKOTKDYTKCVLFQGNMIAYFLSKKICTSPPRDGGYFLSSDSGWKTNFL 240
QY 241 KEGERHLVSKLYTDDMRPEVETVTTTAQETKRGRIQTKKEVSIKTTLKELVHKRVTSPE 300
DB 241 KEGERHLVSKLYTDEMKEPEVETVTTTAQEAKEGRIQTRKEVSIKTTLKELVHKRVTSPE 300
QY 301 DNMMPQDSYIENMAOPGGENLLKNTLEICTLTLARTKTAFDLILEKAETSKLNTFSLPD 360
DB 301 DNMMPQDSYIENMAOPGGENLLKNTLEICTLTLARTKTAFDLILEKAETSKLANFSMAS 360
QY 361 TRTCRIFAFHGWNVYKCHAI CCVLRNQGKRNALVPHGPASTGKSTIIAQAIQAQVGNVG 420
DB 361 TRTCRIFAEHGWNIYKCHAI CCVLRNQGKRNVLVPHGPASTGKSTIIAQAIQAQVGNVG 420
QY 421 CYNAANVPFPNDCTNKNLIWVEEAGNFQGVNQFKAI CSQTIRIDQKGGSKQIEPTP 480
DB 421 CYNAANVPFPNDCTNKNLIWVEEAGNFQGVNQFKAI CSQTIRIDQKGGSKQIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
DB 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
QY 541 NGYQSTWASYCAKWKVPDMSENWAEKVPPTPINLLGSARSPTTPKSTPLSONYALTPL 600
DB 541 NGYQSTWASYCAKWKVPDMSENWAEKVPPTPINLLGSARSPTTPKSTPLSONYALTPL 600
QY 601 ASDLEDLALPEWSTPNTPVAGTAETQNTGEGAGKACQDQGLSPTWSEIEDLACFAGPEP 660
DB 601 ASDLADLALPEWSTPNTPVAGTAASQNTGEGAGTACQAGORSPTWSEIEDLACFASQEQ 660
QY 661 LKDFSEPLNLD 672
DB 661 LKDFSDSLTLD 672

RESULT 7
Q8JV28 PRELIMINARY; PRT; 672 AA.
ID Q8JV28
AC Q8JV28
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
```

```
|:||||:|
661 LKXDFSEPLNLD 672
|:||||:|
661 LESDFNEELTLD 672

RESULT 8
P88899 PRELIMINARY; PRT; 672 AA.
AC P88899;
DT 01-MAY-1997 (Tremblrel. 03, Created)
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Non-capsid protein.
GN Name=NS1;
OS Kilham rat virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=12441;
RN [1]
RP SEQUENCE FROM N.A.
RA Brown D.W., Like A.A.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U79033; AAB38326.1; -.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR InterPro; IPR001257; Parvo_Ns1.
DR Pfam; PF01057; Parvo_Ns1; 1.
SQ SEQUENCE 672 AA; 76056 MW; 0299B36871A0A10A CRC64;

Query Match 91.7%; Score 3335; DB 2; Length 672;
Best Local Similarity 91.4%; Pred. No. 7.1e-238;
Matches 61; Conservative 22; Mismatches 36; Indels 0; Gaps 0;

Qy 1 MAGNAYSDEVLGATNWLKESNQEVFVFKNENVLNGKDIGWNSYKKELODELKSLQ 60
|:||||:|
Db 1 MAGNAYSDEVLGATNWLKSSQEVFVFKNENVLNGKDIGWNSYKKELODELKSLQ 60

Qy 61 RGAETTWQSEDMEWETTVDEMTHKQVFI FDSL VKKCLFEVLNTKNI PGD VNVFVQHEW 120
|:||||:|
Db 61 RGAETTWQSEDMEWESA VDDMTKQVFI FDSL VKKCLFEVLNTKNI PGD VNVFVQHEW 120

Qy 121 GKQGWCHVLIGGKDFSOAQGWRRQLNYSRWLVLTACNVLTAPAERIKLRETAEDN 180
|:||||:|
Db 121 GKQGWCHVLIGGKDFSOQGWRRQLNYSRWLVLTACNVLTAPAERIKLRETAEDS 180

Qy 181 EWTLLTYKHKTKDYTKCVLFGNMIAYFLTKKKISTSPRDGGYFLSSDGSWKTNFL 240
|:||||:|
Db 181 EWTLLTYKHKTKDYTKCVLFGNMIAYFLSKKICTSPRDGGYFLSSDGSWKTNFL 240

Qy 241 KEGERHLVSKLYTDDMPETVETTTAQTAKGRIQTKKEVSIKTTLKELVHKRVTSPE 300
|:||||:|
Db 241 KEGERHLVSKLYTDEMKEPETVETTTAQAEGRIQTKKEVSIKTTLKELVHKRVTSPE 300

Qy 301 DWMMPQDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360
|:||||:|
Db 301 DWMMPQDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLANFSMAS 360

Qy 361 TRTCRI FAFHGWNYKVCHAI CCVLNRQGGKRNVLPHGPASTGKSI IAQAIAQVGNVG 420
|:||||:|
Db 361 TRTCRI FAFHGWNYKVCHAI CCVLNRQGGKRNVLPHGPASTGKSI IAQAIAQVGNVG 420

Qy 421 CYNAA NVFPNDCTNKNLIWVEEAGNFGQVNOFKAI CSQTIRIDQKKGSKQIEPTP 480
|:||||:|
Db 421 CYNAA NVFPNDCTNKNLIWVEEAGNFGQVNOFKAI CSQTIRIDQKKGSKQIEPTP 480

Qy 481 VIMTNNIITVIRIGCEERPEHTQPIRDRMLNTHLTHLPDGLVDKNWPMICAWLVK 540
|:||||:|
Db 481 VIMTNNIITVIRIGCEERPEHTQPIRDRMLNTHLTHLPDGLVDKNWPMICAWLVK 540

Qy 541 NGYQSTWACYCAKWKVPDWSNENWAEKVPPTINLGSASPSPTPKSTPLSNQYALTPL 600
|:||||:|
Db 541 NGYQSTWACYCAKWKVPDWSNENWAEKVPPTINLGSASPSPTPKSTPLSNQYALTPL 600

Qy 601 ASLEDIALEPWSTPNTPVAGTAETONTGEGSKACQDQGLSPTWSEI EEDLACRGAEP 660
|:||||:|
Db 601 ASLEDIALEPWSTPNTPVAGTAETONTGEGSKACQDQGLSPTWSEI EEDLACRGAEP 660

Qy 661 ASLDLADLALFEPWSTPNTPVAGTAETONTGEGSKACQDQGLSPTWSEI EEDLACRGAEP 660
|:||||:|
Db 661 ASLDLADLALFEPWSTPNTPVAGTAETONTGEGSKACQDQGLSPTWSEI EEDLACRGAEP 660

RESULT 9
Q8JV18 PRELIMINARY; PRT; 672 AA.
AC Q8JV18;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Nonstructural protein 1.
GN Name=NS1;
OS Rat minute virus 1a.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=172385;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=22120170; PubMed=12124471;
RA Wan C.H., Soderlund-Venemo M., Pintel D.J., Riley L.K.;
RT "Molecular characterization of three newly recognized rat
RT parvoviruses.";
RL J. Gen. Virol. 83:2075-2083 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Wan C.-H., Pintel D.J., Soderlund-Venemo M., Riley L.K.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF332882; AAM93275.1; -.
DR GO; GO:0019012; C: virion; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR InterPro; IPR001257; Parvo_Ns1.
DR Pfam; PF01057; Parvo_Ns1; 1.
KW Nonstructural protein.
SQ SEQUENCE 672 AA; 76059 MW; 63DB89BBF99E07B3 CRC64;

Query Match 91.5%; Score 3327; DB 2; Length 672;
Best Local Similarity 91.2%; Pred. No. 2.8e-237;
Matches 61; Conservative 23; Mismatches 36; Indels 0; Gaps 0;

Qy 1 MAGNAYSDEVLGATNWLKESNQEVFVFKNENVLNGKDIGWNSYKKELODELKSLQ 60
|:||||:|
Db 1 MAGNAYSDEVLGATNWLKSSQEVFVFKNENVLNGKDIGWNSYKKELODELKSLQ 60

Qy 61 RGAETTWQSEDMEWETTVDEMTHKQVFI FDSL VKKCLFEVLNTKNI PGD VNVFVQHEW 120
|:||||:|
Db 61 RGAETTWQSEDMEWESA VDDMTKQVFI FDSL VKKCLFEVLNTKNI PGD VNVFVQHEW 120

Qy 121 GKQGWCHVLIGGKDFSOAQGWRRQLNYSRWLVLTACNVLTAPAERIKLRETAEDN 180
|:||||:|
Db 121 GKQGWCHVLIGGKDFSOQGWRRQLNYSRWLVLTACNVLTAPAERIKLRETAEDS 180

Qy 181 EWTLLTYKHKTKDYTKCVLFGNMIAYFLTKKKISTSPRDGGYFLSSDGSWKTNFL 240
|:||||:|
Db 181 EWTLLTYKHKTKDYTKCVLFGNMIAYFLSKKICTSPRDGGYFLSSDGSWKTNFL 240

Qy 241 KEGERHLVSKLYTDDMPETVETTTAQTAKGRIQTKKEVSIKTTLKELVHKRVTSPE 300
|:||||:|
Db 241 KEGERHLVSKLYTDEMKEPETVETTTAQAEGRIQTKKEVSIKTTLKELVHKRVTSPE 300

Qy 301 DWMMPQDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360
|:||||:|
Db 301 DWMMPQDSYIEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLANFSMAN 360

Qy 361 TRTCRI FAFHGWNYKVCHAI CCVLNRQGGKRNVLPHGPASTGKSI IAQAIAQVGNVG 420
|:||||:|
Db 361 TRTCRI FAFHGWNYKVCHAI CCVLNRQGGKRNVLPHGPASTGKSI IAQAIAQVGNVG 420

Qy 421 CYNAA NVFPNDCTNKNLIWVEEAGNFGQVNOFKAI CSQTIRIDQKKGSKQIEPTP 480
|:||||:|
Db 421 CYNAA NVFPNDCTNKNLIWVEEAGNFGQVNOFKAI CSQTIRIDQKKGSKQIEPTP 480
```

```
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLYK 540
DB 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLYK 540
QY 541 NGQSTMASYCAKWKGVDPWSENWAPKVPPTPINLLGSARSPTTPKSTPLSQNYALTPL 600
DB 541 NGQSTMACYCAKWKGVDPWSEDAWPKLETPINSIGSMRSPSLTPRSTPLSQNYALTPL 600
QY 601 ASDLELALPEWSTPNTVPVAGTAEQTONTGEGAGSKACODGQLSPTWSEIBEDLRACFGAP 660
DB 601 ASDLADLALPEWSTPNTVPVAGTAASTONTGEGAGSTACQAGRSPTWSEIBEDLRACFSQEQ 660
QY 661 LKQDFSEPLNLD 672
DB 661 LKQDFSDSLTLD 672

RESULT 10
VNC5_PAVH STANDARD; PRT; 672 AA.
AC P03133;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1) (NCVP1).
GN Names=NS1;
OS Hamster parvovirus HI..
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10799;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83112183; PubMed=68232009;
RA Rhode S.L. III, Paradiso P.R.;
RT "Parvovirus genome: nucleotide sequence of H-1 and mapping of its
RL J. Virol. 45:173-184(1983).
CC -!- FUNCTION: Seems necessary for viral DNA replication.
CC -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01457; CA035689.1; -.
DR F03695; U0P011.
DR InterPro; IPR001257; Parvo_NS1.
DR Pfam; PF01057; Parvo_NS1; 1.
KW ATP-binding; DNA replication; Noncapsid protein;
FT NP_BIND 399 406 ATP (Potential).
SQ SEQUENCE 672 AA; 75993 MW; 12F331142F72AA6D CRC64;

Query Match 91.4%; Score 3326; DB 1; Length 672;
Best Local Similarity 91.1%; Pred. No. 3.3e-237;
Matches 612; Conservative 24; Mismatches 36; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVLGATNWLKESKNOEVSFVFKENVOLNGKIDGWSYKKEQLDELKSLQ 60
DB 1 MAGNAYSDEVLGATNWLKESKNOEVSFVFKENVOLNGKIDGWSYKKEQLDELKSLQ 60
QY 61 RGAETTTWDSQSEDEWETTTDEMVKQVIFDLSLVKKCLFEVLTNKNIFPCDVNWFVQHEW 120
DB 61 RGAETTTWDSQSEDEWESAVDDMTKKQVIFDLSLVKKCLFEVLTNKNIFPCDVNWFVQHEW 120
QY 121 GKQDGHCHVLIGKDFPSQAGKWRRLQNLVYNSRWLVACNVQLTPAERIKLREIAEDN 180
DB 121 GKQDGHCHVLIGKDFPSQAGKWRRLQNLVYNSRWLVACNVQLTPAERIKLREIAEDN 180
QY 181 EWTLLTYKHKQTKDYTKCVLFGNNIAYYFLTKKIKISTSPPRDGGYFLSSDSGWKTNFL 240
DB 181 EWTLLTYKHKQTKDYTKCVLFGNNIAYYFLTKKIKISTSPPRDGGYFLSSDSGWKTNFL 240

QY 181 EWTLLTYKHKQTKDYTKCVLFGNNIAYYFLTKKIKISTSPPRDGGYFLSSDSGWKTNFL 240
QY 241 KEGERHLVSKLYTDDMRPETVETTTTAQETKRGRIQTKKEYSIKTKLKLHVHKVTSPE 300
DB 241 KEGERHLVSKLYTDDMRPETVETTTTAQETKRGRIQTKKEYSIKTKLKLHVHKVTSPE 300
QY 301 DNMMPQDSYIEEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360
DB 301 DNMMPQDSYIEEMMAQPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTNFSLPD 360
QY 361 TRICRIFAEGHGWNYKCHAI CCVLNRQGGKRNALVPHGPASTGKSI IAQAIQAAGVNVG 420
DB 361 TRICRIFAEGHGWNYKCHAI CCVLNRQGGKRNALVPHGPASTGKSI IAQAIQAAGVNVG 420
QY 421 CYNAANVNFNDCTNKNLIWVEEAGNFGQVQVNFKAICSGQTIRIDQKGGSKQIEPTP 480
DB 421 CYNAANVNFNDCTNKNLIWVEEAGNFGQVQVNFKAICSGQTIRIDQKGGSKQIEPTP 480
QY 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLYK 540
DB 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNEWPMICAWLYK 540
QY 541 NGQSTMASYCAKWKGVDPWSENWAPKVPPTPINLLGSARSPTTPKSTPLSQNYALTPL 600
DB 541 NGQSTMACYCAKWKGVDPWSEDAWPKLETPINSIGSMRSPSLTPRSTPLSQNYALTPL 600
QY 601 ASDLELALPEWSTPNTVPVAGTAEQTONTGEGAGSKACODGQLSPTWSEIBEDLRACFGAP 660
DB 601 ASDLADLALPEWSTPNTVPVAGTAASTONTGEGAGSTACQAGRSPTWSEIBEDLRACFSQEQ 660
QY 661 LKQDFSEPLNLD 672
DB 661 LKQDFNEELTLD 672

RESULT 11
Q8JUV16 PRELIMINARY; PRT; 672 AA.
ID Q8JUV16;
AC Q8JUV16;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nonstructural protein 1.
GN Names=NS1;
OS Rat minute virus 1b.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=172386;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22120170; PubMed=12124471;
RA Wan C.H., Soderlund-Venermo M., Pintel D.J., Riley L.K.;
RT "Molecular characterization of three newly recognized rat
RT parvoviruses."
RN J. Gen. Virol. 83:2075-2083(2002).
RP SEQUENCE FROM N.A.
RA Wan C.H., Pintel D.J., Soderlund-Venermo M., Riley L.K.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF332883; AM93277.1; -.
DR GO; GO:0019012; C:virion; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR001257; Parvo_NS1.
DR Pfam; PF01057; Parvo_NS1; 1.
KW Nonstructural protein.
SQ SEQUENCE 672 AA; 76201 MW; C2F1A71F6BF49A6 CRC64;

Query Match 91.4%; Score 3323; DB 2; Length 672;
Best Local Similarity 90.9%; Pred. No. 5.5e-237;
Matches 611; Conservative 24; Mismatches 37; Indels 0; Gaps 0;

QY 1 MAGNAYSDEVLGATNWLKESKNOEVSFVFKENVOLNGKIDGWSYKKEQLDELKSLQ 60
DB 1 MAGNAYSDEVLGATNWLKESKNOEVSFVFKENVOLNGKIDGWSYKKEQLDELKSLQ 60
```

```
Db 1 MAGNAYSDEVLGATNWKDKSOEVFSFVKENENVOLNGKDIGNSYRKELQDELKSLQ 60
Qy 61 RGAETTTWDSQEDMEWETTTVDENTKKOVFI FDSLVKKCLFEVLNTKNI PGDVNWFVQHFW 120
Db 61 RGAETTTWDSQEDMEWESAVDDMTKKOVFI FDSLVKKCLFEVLNSTRNIAPSDVTFVQHFW 120
Qy 121 GKDGQWHCHVLIGGKDFSOAQGWRRQLNLYWSRWLVACNVQLTPAERI KLRRETAEDN 180
Db 121 GKDGQWHCHVLIGGKDFSOAQGWRRQLNLYWSRWLVACNVQLTPAERI KLRRETAEDS 180
Qy 181 EWVTLTYKHKOTKDYTKCVLFGNMIAYYFLTKKISTSPRDRGGYFLSSDSGSKTNFL 240
Db 181 EWVTLTYKHKOTKDYTKCVLFGNMIAYYFLTKKISTSPRDRGGYFLSSDSGSKTNFL 240
Qy 241 KEGERHLVSKLYTDDMRPETVETTTAQTKEGRLOTKEVSIKTKLKVHVKRVTSP 300
Db 241 KEGERHLVSKLYTDEMKPETVETTTAQTKEGRLOTKEVSIKTKLKVHVKRVTSP 300
Qy 301 DNMWMPDSDYIENMAOPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTFNLSLPD 360
Db 301 DNMWMPDSDYIENMAOPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTFNFMAN 360
Qy 361 TRTCRIFAFHGWNYKVCHAI CCVLNRQGGKRNVL FHPGASTGKSI IAAQIAQVGNVG 420
Db 361 TRTCRIFAFHGWNYKVCHAI CCVLNRQGGKRNVL FHPGASTGKSI IAAQIAQVGNVG 420
Qy 421 CYNAAVNFPNDCTNKNLIWVEEAGNFQGVNQFKAI CSGQTIRIDQKGSKQIEPTP 480
Db 421 CYNAAVNFPNDCTNKNLIWVEEAGNFQGVNQFKAI CSGQTIRIDQKGSKQIEPTP 480
Qy 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
Db 481 VIMTTNENITVVKIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
Qy 541 NGYQSTWASVCAKWKVPDSENWAEKVPPTINLIGSARSPTTPKSTPLSONYALTP 600
Db 541 NGYQSTWASVCAKWKVPDSENWAEKVPPTINLIGSARSPTTPKSTPLSONYALTP 600
Qy 601 ASLEDLALEPWSPTNPVAGTAETQNTGEGSKACQDGLSPTWSEIEEDLACFCAEP 660
Db 601 ASLEDLALEPWSPTNPVAGTAETQNTGEGSKACQDGLSPTWSEIEEDLACFCAEP 660
Qy 661 LKXDFSEPLNLD 672
Db 661 LKXDFSEPLNLD 672

RESULT 12
VNC3_PAVL3 STANDARD; PRT; 668 AA.
AC P363I;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, last sequence update)
DT 05-JUL-2004 (Rel. 44, last annotation update)
DE Noncapsid protein NS-1 (Nonstructural protein NS1).
GN Names: NS1;
OS Parvovirus LuIII.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=35339;
RN [1]
RP MEDLINE=93297126; PubMed=8517025;
RA Difford N., Chen K.C., Bates R.C., Lederma M.;
RT "the complete nucleotide sequence of parvovirus LuIII and localization
RT of a unique sequence possibly responsible for its encapsidation
RT pattern.";
RL Virology 192:339-345 (1993).
CC -!- FUNCTION: Seems necessary for viral DNA replication.
CC -!- SIMILARITY: Belongs to the parvoviruses noncapsid protein family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M81888; -; NOT ANNOTATED_CDS.
DR PIR; A44276; A44276.
DR InterPro; IPR001257; Parvo NS1.
DR Pfam; PF01057; Parvo NS1; 1.
KW ATP-binding; DNA replication; Noncapsid protein;
KW Nonstructural protein.
FT NP_BIND 399 406 ATP (Potential).
SQ SEQUENCE 668 AA; 75846 MW; CAB69049F8F86B53 CRC64;

Query Match 91.3%; Score 3322; DB 1; Length 668;
Best Local Similarity 91.4%; Pred. No. 6.5e-237;
Matches 614; Conservative 23; Mismatches 31; Indels 4; Gaps 2;

Qy 1 MAGNAYSDEVLGATNWKDKSOEVFSFVKENENVOLNGKDIGNSYRKELQDELKSLQ 60
Db 1 MAGNAYSDEVLGATNWKDKSOEVFSFVKENENVOLNGKDIGNSYRKELQDELKSLQ 60
Qy 61 RGAETTTWDSQEDMEWETTTVDENTKKOVFI FDSLVKKCLFEVLNTKNI PGDVNWFVQHFW 120
Db 61 RGAETTTWDSQEDMEWESSVDELTKKOVFI FDSLVKKCLFEVLNTKNI PGDVNWFVQHFW 120
Qy 121 GKDGQWHCHVLIGGKDFSOAQGWRRQLNLYWSRWLVACNVQLTPAERI KLRRETAEDN 180
Db 121 GKDGQWHCHVLIGGKDFSOAQGWRRQLNLYWSRWLVACNVQLTPAERI KLRRETAEDQ 180
Qy 181 EWVTLTYKHKOTKDYTKCVLFGNMIAYYFLTKKISTSPRDRGGYFLSSDSGSKTNFL 240
Db 181 EWVTLTYKHKOTKDYTKCVLFGNMIAYYFLTKKISTSPRDRGGYFLSSDSGSKTNFL 240
Qy 241 KEGERHLVSKLYTDDMRPETVETTTAQTKEGRLOTKEVSIKTKLKVHVKRVTSP 300
Db 241 KEGERHLVSKLYTDDMRPETVETTTAQTKEGRLOTKEVSIKTKLKVHVKRVTSP 300
Qy 301 DNMWMPDSDYIENMAOPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTFNLSLPD 360
Db 301 DNMWMPDSDYIENMAOPGGENLLKNTLEICTLTARTKTAFDLILEKAETSKLTFNLLAD 360
Qy 361 TRTCRIFAFHGWNYKVCHAI CCVLNRQGGKRNVL FHPGASTGKSI IAAQIAQVGNVG 420
Db 361 TRTCRIFAFHGWNYKVCHAI CCVLNRQGGKRNVL FHPGASTGKSI IAAQIAQVGNVG 420
Qy 421 CYNAAVNFPNDCTNKNLIWVEEAGNFQGVNQFKAI CSGQTIRIDQKGSKQIEPTP 480
Db 421 CYNAAVNFPNDCTNKNLIWVEEAGNFQGVNQFKAI CSGQTIRIDQKGSKQIEPTP 480
Qy 481 VIMTTNENITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
Db 481 VIMTTNENITVVKIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVK 540
Qy 541 NGYQSTWASVCAKWKVPDSENWAEKVPPTINLIGSARSPTTPKSTPLSONYALTP 600
Db 541 NGYQSTWASVCAKWKVPDSENWAEKVPPTINLIGSARSPTTPKSTPLSONYALTP 600
Qy 601 ASLEDLALEPWSPTNPVAGTAETQNTGEGSKACQDGLSPTWSEIEEDLACFCAEP 660
Db 598 -SLEDLALEPWSPTNPVAGTAETQNTGEGSKACQDGLSPTWSEIEEDLACFCAEP 656
Qy 661 LKXDFSEPLNLD 672
Db 657 WKSDSSEQLPMLD 668

RESULT 13
O71159
ID O71159 PRELIMINARY; PRT; 665 AA.
AC O71159;
DT 01-AUG-1998 (T-EMBLrel. 07; Created)
DT 01-AUG-1998 (T-EMBLrel. 07; Last sequence update)
```

01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE Nonstructural protein (Fragment).
 GN Name=NSI;
 OS Kilham rat virus.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
 OX NCBI_TaxID=12441;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=U-Mass;
 RX MEDLINE=98184569; PubMed=9525656;
 RA Ball-Goodrich L.J., Leland S.E., Johnson E.A., Paturzo F.X.,
 RA Jacoby R.O.;
 RT "Rat parvovirus type 1: the prototype for a new rodent parvovirus
 serogroup.";
 RL J. Virol. 72:3289-3299 (1998).
 DR EMBL; AF036711; AAC40695.1; -.
 DR GO; GO:0019012; C:virion; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR InterPro; IPR001257; Parvo_NSI.
 DR Pfam; PF01057; Parvo_NSI; 1.
 DR Nonstructural protein.
 FT NON_TER 1
 SQ SEQUENCE 665 AA; 75375 MW; 778E29043417E409 CRC64;
 Query Match 90.8%; Score 3303; DB 2; Length 665;
 Best Local Similarity 91.4%; Pred. No. 1.6e-235;
 Matches 608; Conservative 22; Mismatches 35; Indels 0; Gaps 0;
 QY 8 DEVLGATNLKESNQEVSVFVKENVQLNGKDIWNYSYKKELOEDLSLQGAETTW 67
 DB 1 DEVLGATNLKDSQEVSVFVKENVQLNGKDIWNYSYKKELOEDLSLQGAETTW 60
 QY 68 DQSEDMEWETTVDENTKKQVFI FDSL VKKCLFEVLNTKNI FPGD VNVVQHEWGDQGW 127
 DB 61 DQSEDMESAVDDMTKKQVFI FDSL VKKCLFEVLNTKNI FPGD VNVVQHEWGDQGW 120
 QY 128 CHVLIGGKDFSQAGKWRRLQNLVYWSRWLVTA CNVLT PPAERIKLREIAEDNEWVTLT 187
 DB 121 CHVLIGGKDFSQAGKWRRLQNLVYWSRWLVTA CNVLT PPAERIKLREIAEDNEWVTLT 180
 QY 188 YKHQTKKDYTKCVLFGNMIAYFLTKKISTSPRDGGYFLSSDSGKWNFLKEGSRHL 247
 DB 181 YKHQTKKDYTKCVLFGNMIAYFLTKKISTSPRDGGYFLSSDSGKWNFLKEGSRHL 240
 QY 248 VSKLYTDDMRPETVETVTTAQTGKRIQTKKVS IKTTLKELVHKVTSPEDDMMWQP 307
 DB 241 VSKLYTDEMKPETVETVTTAQTGKRIQTKKVS IKTTLKELVHKVTSPEDDMMWQP 300
 QY 308 DSYIEMMAQPGGENLLKNTLEICTLT LARTKTAFDLILEKAETSKLTNFSLPDTRICRF 367
 DB 301 DSYIEMMAQPGGENLLKNTLEICTLT LARTKTAFDLILEKAETSKLTNFSLPDTRICRF 360
 QY 368 AFHGWNVYKCHAI CCVLNRQGGKRNVL FHPGASTGKSI IAAIAQAVNGVGCYNAV 427
 DB 361 AEHGWNVYKCHAI CCVLNRQGGKRNVL FHPGASTGKSI IAAIAQAVNGVGCYNAV 420
 QY 428 NPPNDCTNKLILWEEAGNFGQVNOQFKAICSGQTIRIDQKGSKQIEPTPVIMTNE 487
 DB 421 NPPNDCTNKLILWEEAGNFGQVNOQFKAICSGQTIRIDQKGSKQIEPTPVIMTNE 480
 QY 488 NITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVKYGQSTM 547
 DB 481 NITVVRIGCEERPEHTQPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWLVKYGQSTM 540
 QY 548 ASYCAKWKVPDMSENWAEKVPPTPINLLGSARSPTTPKSTPLSQNYALTPLASDL 607
 DB 541 ACYCAKWKVPDMSENWAEKVPPTPINLLGSARSPTTPKSTPLSQNYALTPLASDL 600
 QY 608 ALSPWSTNTPVAGTAETQNTCEAGSKACQDQLSPWSEIEEDLRACFCGAEPKDKPSE 667
 DB 601 ALSPWSTNTPVAGTAETQNTCEAGSKACQDQLSPWSEIEEDLRACFCGAEPKDKPSE 660
 QY 668 PLNLD 672

```
Db 541 NGYOSTWASYSHHWGKVPDSENWAEPAWPTINSIGSARSTQSTATSLQSNYALTPL 600
Qy 601 ASDLEDLALPWSPTNPVAGTAETQNTGEAGSKACODGQLSPWSEIEDLRACFCAEP 660
Db 601 ASDLADLALPWSPTNPVAFVPTGSNT-NTGGRNSQTARSPWSEIADLRACFSOEQ 659
Qy 661 LKKDFSEPLNLD 672
Db 660 LESDFNEELTLD 671

RESULT 15
P89513 ID P89513 PRELIMINARY; PRT; 668 AA.
AC P89513;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nonstructural protein 1.
OS Feline panleukopenia virus (FPV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10786;
RN [1]
RP SEQUENCE FROM N.A.
RA Horiuchi M.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB000053; BAA19014.1; -.
DR GO: GO:0019012; C:virion; IEA.
DR GO: GO:0000166; F:nucleotide binding; IEA.
DR GO: GO:0019079; F:viral genome replication; IEA.
DR InterPro: IPR003593; AAA ATPase.
DR InterPro: IPR001257; Parvo NS1.
DR Pfam: PF01057; Parvo NS1; I.
DR SMART: SMO0382; AAA; 1.
KW Nonstructural protein.
SQ SEQUENCE 668 AA; 76743 MW; 84A210104D9D5788 CRC64;
```

```
Query Match 73.0%; Score 2655; DB 2; Length 668;
Best Local Similarity 73.4%; Pred. No. 1.5e-187;
Matches 493; Conservative 70; Mismatches 103; Indels 6; Gaps 5;

Qy 1 MAGNAYSDEVLGATNMLKEKSNQEVFSFVKPNENVLQNGKDIGWNSYKKELOEDELKSLQ 60
Db 1 MSGNQYTEEVWEGVNLKHAENAEFSFVKCDNVQLNGKDVRRNNYTKPIQNEELTSLI 60

Qy 61 RGAETTWQDS--EDMEWETTVDEMTKKQVFIQSLVKKCLFEVLNTKNIKPPGDVNVFVQH 118
Db 61 RGAQTAMDQTEEBEMDWESEVDSLAKKQVQTFDALIKKCLFEVFSVKNIPEKCVWFIOH 120

Qy 119 EWGKDQGHCHVLIGGKDFGQAGKWRRLQVNVYWSRWLVTAQNVOLTPAERIKLREIAE 178
Db 121 EWGKDQGHCHVLLSHSKNLQATGKWLRRQNNYWSRWLVLCVNLTPTEKIKLREIAE 180

Qy 179 DNEWVTLTYKHQTKDYTKCVLFQGNMIAYFLTKKISTSPRDGGYFLSSDSGWKTN 238
Db 181 DSEWVTLTYRHQTKDYVYKVVHFGNMIAYFLTKKI-VHMTKESGYFLSTDGSKFN 239

Qy 239 FLKEGERHLSKLYTDDMRPETVETTTTAQETKRGRIQTKKEVSIKTKILKELVHKRVTS 298
Db 240 FMKHQDRHTVSTLYTEQKPEVETTTTAQETKRGRIQTKKEVSIKTLRLDLVSKRVTS 299

Qy 299 PEDWMMMQPDSYIEMWAPQGGENLLKNTLEICTLTARTKTAFLDILEKAETSKLTNFSL 358
Db 300 PEDWMMMQPDSYIEMWAPQGGENLLKNTLEICTLTARTKTAFLDILEKADNTKLTNFDL 359

Qy 359 PDTRTCRIFAFHGNVYKVCIAICCVLNROGGRNVAFLFHGPASTGKSIQAIAQAVGN 418
Db 360 ANSRTQCFRMHGNWIKVCHAIACVLNRQGRKNTVLFHGPASTGKSIQAIAQAVGN 419

Qy 419 VGCYNAANVNFENDCTNKNLIWVEAGNFGQVNOFKALCSGQTRIDQKKGSKOIEP 478
Db 420 VGCYNAANVNFENDCTNKNLIWIEAGNFGQVNOFKALCSGQTRIDQKKGSKOIEP 479
```

```
Qy 479 TPVIMTTNENITVVRIGCEERPEHTOPIRDRMLNIHLTHLPGDFGLVDKNWPMICAWL 538
Db 480 TPVIMTTNENITVVRIGCEERPEHTOPIRDRMLNIKLKVKCLPGDFGLVDKEWPLICAWL 539
Qy 539 VKNGYOSTWASYCAKWKGVPDWSENWAEKVPPTPINLGLSARSPTFTPKSTPLSQNYALT 598
Db 540 VKHGYOSTWANYTHHWGKVPEDENWAEPKIQEGINSPG-CKOLETQAAASNPQSDHVL 598
Qy 599 PLASDLEDLALPWSPTNPVAGTAETQNTGEAGSKACODGQLSPWSEIEDLRACFCA 658
Db 599 PLTPDVVDLALPWSPTNPVAGTAETQNTGEAGSKACODGQLSPWSEIEDLRACFCA 656
Qy 659 EPLKXDFSEPLN 670
Db 657 EQLEEDFRDLD 668
```

Search completed: January 22, 2005, 03:40:47
Job time : 108.333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2005, 23:46:57 ; Search time 8541.75 Seconds
(without alignments)
11177.792 Million cell updates/sec

Title: US-10-069-056-16
Perfect score: 2019
Sequence: 1 atggctggaatgcttactc.....agccgctgaactggactaa 2019

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_cm.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2019	100.0	2019	6	AX137751 Sequence
2	2017.4	99.9	2019	6	AX137736 Sequence
3	2017.4	99.9	5081	14	V01115 Minute viru
4	2017.4	99.9	5149	14	J02275 Minute viru
5	2015.8	99.8	2019	6	AX137739 Sequence
6	2015.8	99.8	2019	6	AX137743 Sequence
7	2015.8	99.8	2019	6	AX137747 Sequence
8	1905.4	94.4	5085	14	MV01032 Minute viru
9	1902.2	94.2	4764	14	MU034253
10	1899	94.1	5144	14	MPU12469
11	1899	94.1	5144	14	MU034254
12	1873	92.8	4764	14	MU034255
13	1862.2	92.2	4773	14	MU034256
14	1855.8	91.9	4761	14	MU034256
15	1771	87.7	5121	6	CQ786765 Sequence
16	1662.2	82.3	4904	14	AF321230
17	1659	82.2	4795	14	AF332884
18	1657.4	82.1	4927	14	KRU079033
19	1656.6	82.1	5135	14	PVRSEQ

20	1655.8	82.0	4816	14	AF332883	AF332883 Rat minut
21	1652.6	81.9	4813	14	AF332882	AF332882 Rat minut
22	1644.6	81.5	5176	14	PARH1	X01457 Parvovirus
23	1644.4	81.4	3995	14	AF036711	AF036711 Kilham ra
24	1275	63.2	4936	14	AF036710	AF036710 Rat parvo
25	1051.6	52.1	2007	14	AB000062	AB000062 Feline pa
26	1050	52.0	2007	14	AB000048	AB000048 Feline pa
27	1050	52.0	2007	14	AB000057	AB000057 Feline pa
28	1050	52.0	5124	14	PVFPVC	M3246 Feline panl
29	1046.8	51.8	2007	14	AB000049	AB000049 Feline pa
30	1046.8	51.8	2007	14	AB000058	AB000058 Feline pa
31	1046.8	51.8	2007	14	AB000060	AB000060 Feline pa
32	1046.8	51.8	2007	14	AB000069	AB000069 Feline pa
33	1045.2	51.8	2007	14	AB000053	AB000053 Feline pa
34	1045.2	51.8	2007	14	AB000063	AB000063 Feline pa
35	1045.2	51.8	2007	14	AB000065	AB000065 Feline pa
36	1045.2	51.8	2007	14	AB000067	AB000067 Feline pa
37	1045.2	51.8	5075	14	PVCY1A	D26079 Canine parv
38	1043.6	51.7	2007	14	AB000051	AB000051 Feline pa
39	1043.6	51.7	4983	14	PAPVNS1	X58115 Feline panl
40	1043.6	51.7	5124	14	PVCPVC	M3245 Canine parv
41	1042	51.6	2007	14	AB000055	AB000055 Feline pa
42	1042	51.6	5323	14	PVCCPN	M19296 Canine parv
43	1040.4	51.5	5049	6	AR043629	AR043629 Sequence
44	1040.4	51.5	5049	6	AR043630	AR043630 Sequence
45	1040.4	51.5	5094	14	PVPRPD	D00765 Mink enteri

ALIGNMENTS

RESULT 1
AX137751
LOCUS Sequence 16 from Patent EP1077260.
DEFINITION AX137751 linear PAT 30-MAY-2001
ACCESSION AX137751
VERSION AX137751.1 GI:14273925
KEYWORDS
SOURCE Mice minute virus
ORGANISM Mice minute virus
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE 1
AUTHORS Nuesch, J. and Rommelaere, J.
TITLE Parvovirus ns1 variants
JOURNAL Patent: EP 1077260-A 16 21-FEB-2001;
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts
(DE)

FEATURES
source Location/Qualifiers
1..2019
/organism="Mice minute virus"
/mol_type="unassigned DNA"
/db_xref="taxon:10794"
/note="Parvovirus NS1 variant"

CDS
1..2019
/notes="unnamed protein product"

/codon_start=1
/protein_id="CAC39997.1"
/db_xref="GI:14273926"
/translation="MAGNAYSDEVLGATNWLKESNQEVFSFVKHNVQNGKIDIGW
NSYKELQEDLQSLQGAETTWDSQEDMEWETTVDEMTKKOVFIQDSLVKKCLFEVL
QTKKEVSIKTLKELVKRVTSPEDMMQDPSYIEMMAQPGENLLKNTLEICTLTL
ARTKAFDLILEKETSNTSLPDRTRCFIAPFGWNVKVCIAICVVEAGNFGQR
NTKNIIPGDVNVFQHEWKGQDGHVLIIGKDSQAQGWKRWLVYWSPLWTA
CNVQTPPERRIKLREIAEDNFWTLTYKHQTKDYTKCVLFGNNIAYFLTKKIS
TSPPDGGYFLSDSGKTNFGEHGVLSKLYTDMPETVETVTTVYATKPKRKI
NTVLFHGPASTGKSIQAIAQAVGNVCYNAANVNFNFNDCTNKLIVVEAGNFGQR
QYNQKACSGAIRIDQKSKSQIEPTFPVIMTNENTIVRIGCEEPHTQPIRD
RMLNHLHTLPDGLGVNDWPMICAWLVKNGVQSTWASCAKKGKVPDSENWAE
PKVPPIINLGSARSPTTPKSTPLSQNTALTPGLASDLEDLALSPWSTNTVPVGTAE
TQNTGEAGSKACQDQQLSPTWSEIEDLACFGAEPLKKDFSEPLND"

ORIGIN

Query Match 100.0%; Score 2019; DB 6; Length 2019;

Best Local Similarity 100.0%; Pred. No. 0; Matches 2019; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	ATGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA	60
Db	1	ATGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA	60
Qy	61	AGTAACGAGGAAGTGTCTCATTTTGTGTTTAAATAAGAAATGTTCAACTGAAATGGAAAA	120
Db	61	AGTAACGAGGAAGTGTCTCATTTTGTGTTTAAATAAGAAATGTTCAACTGAAATGGAAAA	120
Qy	121	GATATCGGATGGAATAGTTTACAAAAAGAGCTGAGAGGACGAGCTGAAATCTTTACAA	180
Db	121	GATATCGGATGGAATAGTTTACAAAAAGAGCTGAGAGGAGCAGCTGAAATCTTTACAA	180
Qy	181	CGAGGAGCGGAATCTCTGGGACCAAGCGAGGACATGAAATGGGAAACACAGTGGAT	240
Db	181	CGAGGAGCGGAATCTCTGGGACCAAGCGAGGACATGAAATGGGAAACACAGTGGAT	240
Qy	241	GAAATGACCAAAAGCAAGTATTCATTTTGTGTTTAAATAAGAAATGTTTATTGAA	300
Db	241	GAAATGACCAAAAGCAAGTATTCATTTTGTGTTTAAATAAGAAATGTTTATTGAA	300
Qy	301	GTGCTTAACCAAGAATATATTTCCTGGTATGTTTAAATGGTTTGGCAACATGAATGG	360
Db	301	GTGCTTAACCAAGAATATATTTCCTGGTATGTTTAAATGGTTTGGCAACATGAATGG	360
Qy	361	GGAAAGACCAAGCTGGCACTGCCATGTAATTTGTTTAAATAAGAAATGTTTATTGAA	420
Db	361	GGAAAGACCAAGCTGGCACTGCCATGTAATTTGTTTAAATAAGAAATGTTTATTGAA	420
Qy	421	CAAGGGAATGGTGGAGAGCACTAAATGTTTACTTGGAGCAGATGGTTGGTAAACAGCC	480
Db	421	CAAGGGAATGGTGGAGAGCACTAAATGTTTACTTGGAGCAGATGGTTGGTAAACAGCC	480
Qy	481	TGTAATGTGCAACTAACCAACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAACAAT	540
Db	481	TGTAATGTGCAACTAACCAACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAACAAT	540
Qy	541	GAGTGGGTTACTTACTTACTTATAGCATAAGCAACCAAAAGAACTATACCAAGTGT	600
Db	541	GAGTGGGTTACTTACTTACTTATAGCATAAGCAACCAAAAGAACTATACCAAGTGT	600
Qy	601	GTTCCTTTTGGAAACATGATGCTTACTTATTTTAACTAAAAAGAAAAATAAGCACTAGT	660
Db	601	GTTCCTTTTGGAAACATGATGCTTACTTATTTTAACTAAAAAGAAAAATAAGCACTAGT	660
Qy	661	CCACCAAGAGACGAGGCTATTTCTTAGCAGTGAATGCTGGCTGGAAAACTAACTTTTAA	720
Db	661	CCACCAAGAGACGAGGCTATTTCTTAGCAGTGAATGCTGGCTGGAAAACTAACTTTTAA	720
Qy	721	AAAGAGGCGAGCGCCATCTAGTCAGCAAACTATACATGATGATGCGGCCAGAAACG	780
Db	721	AAAGAGGCGAGCGCCATCTAGTCAGCAAACTATACATGATGATGCGGCCAGAAACG	780
Qy	781	GTTGAAACCAACAGTAACCACTGCGCAGGAACTAAGCGCGCAGAAATCAAACTAAAAA	840
Db	781	GTTGAAACCAACAGTAACCACTGCGCAGGAACTAAGCGCGCAGAAATCAAACTAAAAA	840
Qy	841	GAGTTTCTATTAAACTACATTAAGAGCTGGTGATTAAGAGTAACCTCACAGAG	900
Db	841	GAGTTTCTATTAAACTACATTAAGAGCTGGTGATTAAGAGTAACCTCACAGAG	900
Qy	901	GACTGGATGATGTCAGCAGACAGTTATGTAATGATGCTCAACACAGGTGAGAA	960
Db	901	GACTGGATGATGTCAGCAGACAGTTATGTAATGATGCTCAACACAGGTGAGAA	960
Qy	961	AACCTGCTGAAAAATACGCTAGAGATTGTACACTAACTAGCCAGAACCAAAACAGCA	1020
Db	961	AACCTGCTGAAAAATACGCTAGAGATTGTACACTAACTAGCCAGAACCAAAACAGCA	1020
Qy	1021	TTTGACTTAATTTTAGAAAAAGCTGAAACGAGCAAACTAACCTTTTCACTGCTGAC	1080

RESULT 2
AX137736
LOCUS
DEFINITION
ACCESSION
VERSION

AX137736 2019 bp DNA linear PAT 30-MAY-2001
Sequence 1 from Patent EP1077260.
AX137736 GI:14273909

KEYWORDS	Mice minute virus	Db	481	TGTAATGTGCAACTAACACCCAGCTGAAAGAAATTAATACTAAGAGAAATAGCAGAAGCAAT	540
SOURCE	Mice minute virus	Qy	541	GAGTGGTTACTTACTTACTTATATAGCATTAAGCAACCAAAAGACATATACCAAGTGT	600
ORGANISM	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.	Db	541	GAGTGGTTACTTACTTACTTATATAGCATTAAGCAACCAAAAGACATATACCAAGTGT	600
REFERENCE	1	Qy	601	GTTCCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAAAATAAGCACTAGT	660
AUTHORS	Nuesch, J. and Rommelaere, J.	Db	601	GTTCCTTTTGGAAACATGATGCTTACTATTTTAACTAAAAAGAAAAATAAGCACTAGT	660
TITLE	Parvovirus ns1 variants	Qy	661	CCACCAAGAGACGAGGCTATTTCTTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA	720
JOURNAL	Patent: EP 1077260-A 1 21-FEB-2001; Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts (DE)	Db	661	CCACCAAGAGACGAGGCTATTTCTTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA	720
FEATURES	Location/Qualifiers	Qy	721	AAAGAAGGAGCGGCCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG	780
source	1. .2019	Db	721	AAAGAAGGAGCGGCCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG	780
	/organism="Mice minute virus"	Qy	781	GTTGAAACACACAGTAAACCTGCGCAGGAAACTTAAGCGGCGAGAACTTCAAACTAAAAAA	840
	/mol_type="unassigned DNA"	Db	781	GTTGAAACACACAGTAAACCTGCGCAGGAAACTTAAGCGGCGAGAACTTCAAACTAAAAAA	840
	/db_xref="taxon:10794"	Qy	841	GAAAGTTTCTATTAAAACTACACTTAAGAGCTGGTCATATAAAGAGTAACTTCAACAGAG	900
	/note="Wildtype Parvovirus NS1"	Db	841	GAAAGTTTCTATTAAAACTACACTTAAGAGCTGGTCATATAAAGAGTAACTTCAACAGAG	900
	1. .2019	Qy	901	GACTGATGATGATGAGCAGACAGTACATTGAAATGATGCTCAACAGGTGGAGAA	960
	/notes="unnamed protein product"	Db	901	GACTGATGATGATGAGCAGACAGTACATTGAAATGATGCTCAACAGGTGGAGAA	960
	/codon_start=1	Qy	961	AACCTGCTGAAAATACGCTAGAGATTTGTACACTAACTCTACCCAGAACCAAAACAGCA	1020
	/protein_id="CAC39989.1"	Db	961	AACCTGCTGAAAATACGCTAGAGATTTGTACACTAACTCTACCCAGAACCAAAACAGCA	1020
	/db_xref="GI:14273910"	Qy	1021	TTTGACTTAATTTAGAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCTGAC	1080
	/translation="MAGNAYSDVLGATNWLKESNQEVFSFVFKENYVQLNGKIDIG NSYKLEDELKSLQGAETTWQSEDMEWETTVDEMFKQVFIIDSLVKCLFVL NTKNIIPGDVNVFVQHEWKGDKGWHCHVLIGKDFSOAGCKWRROLNVYWSRLVTA CNVQLPAERIKLRETAEDNEWILLTYKHQTKDYTKCVLFGNMIAFYFLTKKIS TSPPRGYFLSSDSGNFNLKEGERHLSKLYIDDMRPETVETITVTAQETKGR QTKVESIKTTLKELVHKRVTSPEDMMWQPDYSIEMMAQPGGENILKNTLEICTLTL ARTKADLTLEAKETSILNFSLPDTTRICRIIFAFHWNVYKVCNAICCVLNKQGGK NTVLFGPASTGKSIITAOIAOAVGNVGCYNAANVFPNDCTNKLIVVEEAGNFGQ OVNQFKAISGOTIRIDOKGSKQLEPTEVIMTNENITVAVIGCEERPEHTQPIRD RMLNIHLTHLPGDFGLVDKNWPMICMLVKNYQSTWASCYCAKNGKVPDWSNWE KPYPTPINLIGSFFLTPKSONYALTPLASDLELALPELWSTNPTNPVAGTAE TQNTGAGSKACQDQLSPTWSIEBIDLACFGAEFLKDFSEPLND"	Db	1021	TTTGACTTAATTTAGAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCTGAC	1080
ORIGIN	Query Match 99.98; Score 2017.4; DB 6; Length 2019; Best Local Similarity 100.0%; Pred. No. 0; Matches 2019; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	Qy	1081	ACAAGAACCTGCAAGATTTTGTGCTTTCATGCTGGAACATATGTTAAAGTTGCGCATGCT	1140
		Db	1081	ACAAGAACCTGCAAGATTTTGTGCTTTCATGCTGGAACATATGTTAAAGTTGCGCATGCT	1140
		Qy	1141	ATTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATACTGTTTATTTTCATGGACCA	1200
		Db	1141	ATTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATACTGTTTATTTTCATGGACCA	1200
		Qy	1201	GCAGACAGGCAAACTCTATTATTCAGAACCATAGCAGCAAGCATGTTGGCAATGTTGGT	1260
		Db	1201	GCAGACAGGCAAACTCTATTATTCAGAACCATAGCAGCAAGCATGTTGGCAATGTTGGT	1260
		Qy	1261	TGCTATAATGCAAGCAATGTAACTTTTCCATTTTAAATGACTGTACCAACCAAGAACTTGATT	1320
		Db	1261	TGCTATAATGCAAGCAATGTAACTTTTCCATTTTAAATGACTGTACCAACCAAGAACTTGATT	1320
		Qy	1321	TGGGTAGAGAGAGCTGGTAACTTTGACAGCAAGTAAACCCAGTTTAAAGCCATTGCTCT	1380
		Db	1321	TGGGTAGAGAGAGCTGGTAACTTTGACAGCAAGTAAACCCAGTTTAAAGCCATTGCTCT	1380
		Qy	1381	GGTCAAGCTATTTCGCTTATGATCAAAAGGAAAGGCAAGAGATTTGAACCAACCA	1440
		Db	1381	GGTCAAGCTATTTCGCTTATGATCAAAAGGAAAGGCAAGAGATTTGAACCAACCA	1440
		Qy	1441	GTGATCATGACCAAAATGAGAACTTTACAGTGGT CAGAA TAGGCTGCGAAGAAAGCA	1500
		Db	1441	GTGATCATGACCAAAATGAGAACTTTACAGTGGT CAGAA TAGGCTGCGAAGAAAGCA	1500
		Qy	1501	GAAACACTCAACCAATCAGAGACAGATTCGTTAACTTCACTTAACACATACCTTGCCT	1560
		Db	1501	GAAACACTCAACCAATCAGAGACAGATTCGTTAACTTCACTTAACACATACCTTGCCT	1560
		Qy	1561	GGTGACTTGGTTGGTTGACAAAAATGATGCCCATGATTTGCTGCTGTTGGTAAAG	1620
		Db	1561	GGTGACTTGGTTGGTTGACAAAAATGATGCCCATGATTTGCTGCTGTTGGTAAAG	1620

Qy	1621	AATGGTTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGCAAAAGTTCTGATGG	1680
Db	1621	AATGGTTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGCAAAAGTTCTGATGG	1680
Qy	1681	TCAGAAACTGGGGCGGAGCCAAAGGTGCCAACTCTTAATAATTACTAGGTTGGCAACG	1740
Db	1681	TCAGAAACTGGGGCGGAGCCAAAGGTGCCAACTCTTAATAATTACTAGGTTGGCAACG	1740
Qy	1741	TCACATTCACGACACCCGAAAGATGACGCTCTCAGCCGAGCAACATATGACCTTCAACTT	1800
Db	1741	TCACATTCACGACACCCGAAAGATGACGCTCTCAGCCGAGCAACATATGACCTTCAACTT	1800
Qy	1801	GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAATACTCTGTGTCG	1860
Db	1801	GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAATACTCTGTGTCG	1860
Qy	1861	GGCACTCGAGAAACCCAGAACACCTGGGGAAGCTGTTCCAAAGCTGCCAAGATGTCAC	1920
Db	1861	GGCACTCGAGAAACCCAGAACACCTGGGGAAGCTGTTCCAAAGCTGCCAAGATGTCAC	1920
Qy	1921	CTGAGCCCAACTGGTTCAGAGATCGAGAGGATTTGAGAGCGTCTCGGTGCGGAACCG	1980
Db	1921	CTGAGCCCAACTGGTTCAGAGATCGAGAGGATTTGAGAGCGTCTCGGTGCGGAACCG	1980
Qy	1981	TTGAAGAAAGACTTCACGAGCGCGCTGAACTTGGACTAA	2019
Db	1981	TTGAAGAAAGACTTCACGAGCGCGCTGAACTTGGACTAA	2019
RESULT 3			
PAMV2			
LOCUS	PAMV2	5081 bp	DNA linear VRL 10-FEB-1999
DEFINITION	Minute virus of mice with two major open reading frames (genome).		
ACCESSION	V0115		
VERSION	V0115.1	GI:60911	
KEYWORDS	coat protein; genome; origin of replication; overlapping genes; terminal repeat.		
SOURCE	Mice minute virus		
ORGANISM	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.		
REFERENCE	1 (bases 1 to 5081)		
AUTHORS	Astell,C.R., Thomson,M., Merchlinsky,M. and Ward,D.C.		
TITLE	The complete DNA sequence of minute virus of mice, an autonomous parvovirus		
JOURNAL	Nucleic Acids Res. 11 (4), 999-1018		(1983)
MEDLINE	83143341		
PUBMED	6298737		
COMMENT	The messenger RNA of this virus (colinear with the strand listed) is spliced, but the exact splice sites are not known. The listed strand is complementary to the one which is included in the virion.		
FEATURES	Location/Qualifiers		
source	1..5081		
	/organism="Mice minute virus"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:10794"		
	114..2279		
	/note="unlabeled protein product; coding sequence"		
	/codon_start=1		
	/protein_id="CAA24309.1"		
	/db_xref="GI:60912"		
	/db_xref="GOA:P03134"		
	/db_xref="Swiss-Prot:P03134"		
	/translation="MISGSGSLNQAKRWKFWKVKYQLLKSVYLPFFHSVSRDAOKE SNQLTWAGNAYSDEVIGATNWLKEKNEQVFSVFKNNVQNLGDKIDGWSYKKELOE DELKSLQRGAEFTWDSQEDMEWETVDEMTKQVIFDSLAVKCLPEVLNKNIFPFG VNFVFEWEGNDQGHVLLGKDFSAQGWKRWQLNVYSLRWLVLTACNQLTPAE RKILRETAENQWTLTYKHKOTKDYTKVLFGNMIAYFLTKKISTSPRDGQY FLSDSGWKTNFLKEGHEHLYSKLYTDDMPETVTTTAQETKEGRICQTKKEVSIK TTIKELVHKVTSPEWMMQPSYIEMMQPGENLLKNLEICITLTARTKATPDL ILBEAKTINFLSPDRTRCTPFAFHMVNVKVAICCLVNRQGGKRNVLTFHPGA STGKSIQAQAVGNVCYNAANVPFNDTNKLIWVEAGNFGQGVNQFKAIC SGOTIRIDQKGGSKQIETPPYTNINETVVRIGCEERPEHTQPIRDRMLNHLTH SGOTIRIDQKGGSKQIETPPYTNINETVVRIGCEERPEHTQPIRDRMLNHLTH		
CDS			

Db 561 GTCTTAAACAAGAAATATNTTCTGGTGTATTTAAATGGTTTGTGCAACATGATGG 620
Qy 361 GGAAGAAGCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAGGACATTTAGTCAAGCT 420
Db 621 GGAAGAAGCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAGGACATTTAGTCAAGCT 680
Qy 421 CAAGGGAATGTTGGAGAGGCACTAATATTTTACTGTGAGCAGATGGTTGTTAAACAGCC 480
Db 681 CAAGGGAATGTTGGAGAGGCACTAATATTTTACTGTGAGCAGATGGTTGTTAAACAGCC 740
Qy 481 TGTAAATGTGCAACTAACAAGGCTGAAAGAAATTTAACTAAGAGAAATAGCAGAGAACAAT 540
Db 741 TGTAAATGTGCAACTAACAAGGCTGAAAGAAATTTAACTAAGAGAAATAGCAGAGAACAAT 800
Qy 541 GAGTGGTTTACTTACTTACTTATAAGCATTAAGCAAAACCAAAAAAGCATATACCAAGTGT 600
Db 801 GAGTGGTTTACTTACTTACTTATAAGCATTAAGCAAAACCAAAAAAGCATATACCAAGTGT 860
Qy 601 GTTCTTTTGGAAACATGATGCTTACTTATTTTAACTTAAAGAAAGAAATTAAGCACTAGT 660
Db 861 GTTCTTTTGGAAACATGATGCTTACTTATTTTAACTTAAAGAAAGAAATTAAGCACTAGT 920
Qy 661 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTGACTCTCGCTGGAATACTAACTTTTAA 720
Db 921 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTGACTCTCGCTGGAATACTAACTTTTAA 980
Qy 721 AAAGAAAGGCGAGGCGCCATCTAGTGAGCAAACTATACACTGATGACATCGCGCCAGAAACG 780
Db 981 AAAGAAAGGCGAGGCGCCATCTAGTGAGCAAACTATACACTGATGACATCGCGCCAGAAACG 1040
Qy 781 GTTGAACCAACAGTAAACCACTGCGCAGGAACTAAGCGCGCAGAAATTTCAAACTAAAAA 840
Db 1041 GTTGAACCAACAGTAAACCACTGCGCAGGAACTAAGCGCGCAGAAATTTCAAACTAAAAA 1100
Qy 841 GAAGTTTCTATTAAAACTACATTAAGAGCTGGTGCATFAAAGAGTAACTTCACTCAGAG 900
Db 1101 GAAGTTTCTATTAAAACTACATTTAAGAGCTGGTGCATFAAAGAGTAACTTCACTCAGAG 1160
Qy 901 GACTGGATGATGATGACGACAGACAGTTACATTTGAAATGATGCTCAACAGGCTGGAGAA 960
Db 1161 GACTGGATGATGATGACGACAGACAGTTACATTTGAAATGATGCTCAACAGGCTGGAGAA 1220
Qy 961 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAATCTAGCCAGAACCAAAACAGCA 1020
Db 1221 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAATCTAGCCAGAACCAAAACAGCA 1280
Qy 1021 TTTGACTTAATTTAGAAAAGCTGAAACAGCAAACTAACCACTTTTCACTGCTGAC 1080
Db 1281 TTTGACTTAATTTAGAAAAGCTGAAACAGCAAACTAACCACTTTTCACTGCTGAC 1340
Qy 1081 ACAAGAACCTGCAAGATTTTGTCTTTTCTGCTGGAATCTATGTTAAAGTTTGCCATGCT 1140
Db 1341 ACAAGAACCTGCAAGATTTTGTCTTTTCTGCTGGAATCTATGTTAAAGTTTGCCATGCT 1400
Qy 1141 ATTTGCTGTGTTTTAAACAGAGGAGGCAAGAAATCTGTTTTTATTTATGAGCA 1200
Db 1401 ATTTGCTGTGTTTTAAACAGAGGAGGCAAGAAATCTGTTTTTATTTATGAGCA 1460
Qy 1201 GCAGACAGGCAAACTATTTATGCAAGCCATAGCAACAGAGTTGCAATGTTGTT 1260
Db 1461 GCAGACAGGCAAACTATTTATGCAAGCCATAGCAACAGAGTTGCAATGTTGTT 1520
Qy 1261 TGCTATAATCCAGCAATGTAATTTTCACTTTTAAAGTGTACCAAGAACTTTGATT 1320
Db 1521 TGCTATAATCCAGCAATGTAATTTTCACTTTTAAAGTGTACCAAGAACTTTGATT 1580
Qy 1321 TGGGTAGAAGAGCTGTAATTTTGAACAGCAAGTAAACAGTTTAAAGCAATTTGCTCT 1380
Db 1581 TGGGTAGAAGAGCTGTAATTTTGAACAGCAAGTAAACAGTTTAAAGCAATTTGCTCT 1640
Qy 1381 GGTCAAGCTATTCGATTTGATCAAAAAGGAAAGGAGCAAAACAGATTTGAACCAACCA 1440

Db 1641 GGTCAAACTATTTCGCATTGATCAAAAAGGAAAGGAGCAAAACAGATTTGAACCAACACCA 1700
Qy 1441 GTCATCATGACACAAATGAGAACATTAACAGTGGTTCAGATAGGCTGCGAAGAAACCA 1500
Db 1701 GTCATCATGACACAAATGAGAACATTAACAGTGGTTCAGATAGGCTGCGAAGAAACCA 1760
Qy 1501 GAAACACACTCAACCAATCAGAGACAGAAATGCTTAACTTATCATCTAACATACATCTTGCCT 1560
Db 1761 GAAACACACTCAACCAATCAGAGACAGAAATGCTTAACTTATCATCTAACATACATCTTGCCT 1820
Qy 1561 GGTGACTTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGTCTTGGTTGGTAAAG 1620
Db 1821 GGTGACTTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGTCTTGGTTGGTAAAG 1880
Qy 1621 AATGTTTACCAATCTACCATGAGCAAGTACTGTCTTAAATGGGGCAAAAGTTCTGATTGG 1680
Db 1881 AATGTTTACCAATCTACCATGAGCAAGTACTGTCTTAAATGGGGCAAAAGTTCTGATTGG 1940
Qy 1681 TCAGAAAACTGGCGGAGCCAAAGGTGCCAACTCTCTATAAAATTTACTAGGTTCCGACCG 1740
Db 1941 TCAGAAAACTGGCGGAGCCAAAGGTGCCAACTCTCTATAAAATTTACTAGGTTCCGACCG 2000
Qy 1741 TCACCAATTCACGACACCGAAAAAGTAGCGCTCTCAGCCAGAACTATGCACTAACCTCCACTT 1800
Db 2001 TCACCAATTCACGACACCGAAAAAGTAGCGCTCTCAGCCAGAACTATGCACTAACCTCCACTT 2060
Qy 1801 GAATCGGATCTCGAGGACCTGCTTTAGAGCCTTTAGAGCCTTTGGAGCACACCAATACTCTCTGTTGG 1860
Db 2061 GCATCGGATCTCGAGGACCTGCTTTAGAGCCTTTAGAGCCTTTGGAGCACACCAATACTCTCTGTTGG 2120
Qy 1861 GGCACCTGCAAGAACCCAGAACACTGCGGAGCTGGTTCCAAAGCCTGCCAAGATGCTCAA 1920
Db 2121 GGCACCTGCAAGAACCCAGAACACTGCGGAGCTGGTTCCAAAGCCTGCCAAGATGCTCAA 2180
Qy 1921 CTGAGGCCCACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTCGTTGCGGAACCG 1980
Db 2181 CTGAGGCCCACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTCGTTGCGGAACCG 2240
Qy 1981 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019
Db 2241 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2279

RESULT 4

LOCUS MVMPG 5149 bp ss-DNA linear VRL 22-MAY-1995
DEFINITION Minute virus of mice, complete genome.
ACCESSION J02275 M12520 M12521 M14704
VERSION J02275.1 GI:332293
KEYWORDS alternative splicing; capsid protein; complete genome; nonstructural protein.
SOURCE Mice minute virus
ORGANISM Mice minute virus
REFERENCE 1 (bases 1 to 5149)
AUTHORS Astell,C.R., Thomson,M., Merchinsky,M. and Ward,D.C.
TITLE The complete DNA sequence of minute virus of mice, an autonomous parvovirus
JOURNAL Nucleic Acids Res. 11 (4), 999-1018 (1983)
MEDLINE 83143341
PUBMED 6298737
REFERENCE 2 (bases 1 to 5149)
AUTHORS Astell,C.R., Gardiner,E.M. and Tattersall,P.
TITLE DNA sequence of the lymphotropic variant of minute virus of mice, VMV(i), and comparison with the DNA sequence of the fibrotropic prototype strain J. Virol. 57 (2), 656-669 (1986)
JOURNAL MEDLINE 86115415
PUBMED 3502703
REFERENCE 3 (sites)
AUTHORS Morgan,W.R. and Ward,D.C.
TITLE Three splicing patterns are used to excise the small intron common to all minute virus of mice RNAs

J. Virol. 60 (3), 1170-1174 (1986) 87061199 3783817 PUBLISHED COMMENT	J. Virol. 60 (3), 1170-1174 (1986) 87061199 3783817 PUBLISHED COMMENT	Original source text: Minute virus of mice (strain MVM(p)), passed in mouse 1 (variant A-9) cells. The parvoviridae family contains two groups that infect mammalian hosts: (i) defective (helper-dependent) adeno-associated viruses, and (ii) autonomous (helper-independent) parvoviruses. MVM is a member of the latter group. Both groups have been demonstrated to package both plus and minus strands (in separate particles) of the ss-DNA genome, though the minus strand is more typically packaged in the latter group. The sequence below corresponds to the plus (+) strand, also referred to as the C-strand. The minus (-) strand is also referred to as the V-strand. The 3' and 5' termini both exhibit the potential for forming stable 'fold-back' hairpins; these sequences appear to play a role in replication [1]. The left and right halves of the genome encode two distinct, but overlapping transcriptional units. The transcripts can be summarized [1] (1 map unit (mu) = 51 bp): R1 (4.8 kb): 4.5 mu - 46 mu; 46+ mu - 95 mu R2 (3.3 kb): 4.5 mu - 10.7 mu; 38 mu - 46 mu; 46+ mu - 95 mu R3 (3.0 kb): 40 mu - 46 mu; 46+ mu - 95 mu R3 is the major transcript. There are two major open reading frames, both on the plus (or C) strand. The left side ORF (261-2279) probably encodes a non-capsid protein of 85 kd; the right side ORF probably encodes the viral capsid proteins, VP1 (or A, 83 kd), VP2 (or B, 64 kd), and VP3 (or C, 61 kd). But because of uncertainties about the precise splice points in the transcripts, the exact starts, stops and (possible) intron boundaries are not known. revision 4804 4870 a-65bp-a in [2]; aa in [1] [2] sites [1]. [3] sites; splice sites. Location/Qualifiers 1. 5149 /organism="Mice minute virus" /mol_type="genomic DNA" /strain="MVM(p)" /db_xref="taxon:10794" /lab_host="mouse 1 (variant A-9) cell" 114..2279 /gene="NS1" 114..2279 /gene="NS1" /note="putative" /codon_start=1 /product="nonstructural protein" /protein_id="AAA67108.1" /db_xref="GI:825481" /translations="MISGSGSLNQAKRWAFKVKYKOLLKSVYLFHFSVSDAOKEL SNOLTMAGNAYSDEVLGATNWLKEKSNQEVFVFKNENVLQKDIGNWSYKKELOE DELKSLQGAETDQSEDEWETTVDKQVFIQDLSLVKCLFVNTKNIFFPD VNMVFWHEGDKQCHVLIGKQFSQAQKWRRLQNVWSRWLVATCNQLTPAE RKLRLAEADNEWTLTKYKQKDYTKCVLFGNMAYFLTKKISIPSPDGGY FLSSDCKNTFLKEBERHLVSKLYTDMRPETVTTTFAQETKGRITQKKEVSIK TTLKELVHKVTSPEDEMMQPDYSIEMMAQPGENLLKNTLICTLTARTKTAEDL ILKSAKTNFLSLPTDTRCFRFAFGHWNVYKVAICCVLRQGGKNTVLPHGPA STGKSIQAIAQAVGNCYNAANVFPDNCNKLWVEEAGNFQQQVNFQKAIK SGQTIIRDQKSGKQIEPTVIMTNENITVVRIGCEERPEHTQPIRDLMLNHLTH TLPGDFGLVDKNEPMICAWLVKNGQSTWASYCAWKGVKPDWSENWAPKVPPTNL LGSARSPFTTPKSTPLSQNYALTPLASDLEDALEPWTSTPNTVAGTAETQNTGAGS KACQDGLSTWSEIEDLRACFGAEPKDKDFSEPLND" 200..>2279 /gene="NS1" 261..2279 /gene="NS1" /codon_start=1 /product="nonstructural protein" /protein_id="AAA67109.1" /db_xref="GI:332295" /translation="MAGNAYSDEVLGATNWLKEKSNQEVFVFKNENVLQKDIGW	NSYKKELOEDELKSLQGAETDQSEDEWETTVDKQVFIQDLSLVKCLFVNL NTKNIFPDGVDNFWHEGDKQCHVLIGKQFSQAQKWRRLQNVWSRWLVTA CNVQLTPAEIRKLRLAEADNEWTLTKYKQKDYTKCVLFGNMAYFLTKKIS TPSPDGGYFLSSDCKNTFLKEBERHLVSKLYTDMRPETVTTTFAQETKGRIT QKKEVSIKTTTLKELVHKVTSPEDEMMQPDYSIEMMAQPGENLLKNTLICTLT ATKTAEDLILKSAKTNFLSLPTDTRCFRFAFGHWNVYKVAICCVLRQGGK NTVLPHGPASTGKSIQAIAQAVGNCYNAANVFPDNCNKLWVEEAGNFQQ QNTQKAIKSGQTIIRDQKSGKQIEPTVIMTNENITVVRIGCEERPEHTQPIR RDLMLNHLTHTLPGDFGLVDKNEPMICAWLVKNGQSTWASYCAWKGVKPDWSENWAE LGSARSPFTTPKSTPLSQNYALTPLASDLEDALEPWTSTPNTVAGTAET TQNTGAGSKACQDGLSTWSEIEDLRACFGAEPKDKDFSEPLND" 2002..2398 /gene="VP" 2002..2280 /gene="VP" /note="major transcription start site" /number=1 2006..2280 /gene="VP" /note="minor transcription start site" /number=1 2009..2280 /gene="VP" /note="minor transcription start site" /number=1 2281..2398 /gene="VP" /note="alternative intron" 2281..2376 /gene="VP" /note="alternative intron" join(2286..2316,2399..4557) /gene="VP1" join(2286..2316,2399..4557) /gene="VP1" /codon_start=1 /protein_id="AAA67111.1" /db_xref="GI:332294" /translation="MAPPAKRAKRWVPPGYKYLPGNSLDQGEPTNPDAKAEHDE AYDOVIKSGKNPVLVPSAADORFIDQTKADQWGGKVGHVFTKRAFAPKLATSE GTSVSRAGKTRPPAYIFINQARAKKLTSSAAQSSQTSMDGTSPQSGNVAHSA RVERAADGPGSGGGGSGVSTGSDNQTFRFLDGNVEITATLRLVHLNMP KSENYCRIRVHNTDTSVKNWAKDAHEQIWPWSLDAACMMVAVDSNNILFYICNT MSQLNLVSLDQEIFNVVLKTVTEQDLGGQAIKIYNNDLTANWAVDNNILFYICNT NSMETIGFYPMKPTIASPYRYFCVDRDLSTVYENQEGTVENHVMGTPKGMSQFTI ENTQOITLLRTGDEFATGYTFDTNSVKLTHTWQNRQLGQPLLTFFPEADTQGL TAQGSRHGTTQMVANVSEAIRTPAQVGFQCPHNDFEASRAGPFAAPKVPADITQGV DKRSGVRYSYKQGHENWASHGAPERYTWDETSFGSGRTKQGFQISAPLVPPP LNGILTANPITGKNDIHFSNVFNSTGPTAFSHPSPVYPOQGIWDKLDLHKPLRH ITAPFVKNNAPGQMLVRLGPLNLTQYDPNGATLSRIVTYGTFFWKGLTMRKLRAN TTWNPVYQVSAEDNGNSYMSVTKLPTATGNNQSVPLITRPVARTY" 2286..2354 /gene="VP" /note="ORF1; putative" /codon_start=1 /product="unknown protein" /protein_id="AAA67110.1" /db_xref="GI:825482" /translation="MAPPAKRAKRGKGLRDGLVGY" 2286..2316 /gene="VP" /number=1 2317..2398 /gene="VP" /note="VP intron (alt.)" 2332..2361 /gene="VP" /note="ORF3; putative" /codon_start=1 /product="unknown protein" /protein_id="AAA67112.1" /db_xref="GI:825483" /translation="MVGWNGINV"
---	---	---	--

CDS	2354..2398	/gene="VP"	861	GTCTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAGAAATAAGCACTAGT	920
	/note="ORF2; putative"		661	CCACCAAGAGACGGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAATACTATTCTTTTA	720
	/codon_start=1		921	CCACCAAGAGACGGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAATACTATTCTTTTA	980
	/product="unknown protein"		721	AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGGGCCAGAAACG	780
	/protein_id="AAA67113.1"		981	AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGGGCCAGAAACG	1040
exon	/db_xref="GI:825484"		781	GTTGAAACCAACAGTAACCACTGCGCAGGAAACTTAAGCGCGGCGAGAAATTCAAACATAAAAA	840
	/translations="MFNYLFYRPIITWF"		1041	GTTGAAACCAACAGTAACCACTGCGCAGGAAACTTAAGCGCGGCGAGAAATTCAAACATAAAAA	1100
	2399..>4557		841	GAAAGTTCTTATTAACACTACCTTAAAGAGCTGGTGCATATAAAGAGTAACCTCACCAGAG	900
	/gene="VP1"		1101	GAAAGTTCTTATTAACACTACCTTAAAGAGCTGGTGCATATAAAGAGTAACCTCACCAGAG	1160
	/number=2		901	GACTGGATGATGATGCGCAGCAGACAGTACATTTGAATGATGCTCAACACGAGTGGAGAA	960
	/gene="VP1"		1161	GACTGGATGATGATGCGCAGCAGCAGTACATTTGAATGATGCTCAACACGAGTGGAGAA	1220
	/notes="VP2"		961	AACCTCTCTGAAATAACGCTAGAGATTTGTACATACTAAGCTCTAGCCAGAAACAAACAGCA	1020
	/codon_start=1		1221	AACCTCTCTGAAATAACGCTAGAGATTTGTACATACTAAGCTCTAGCCAGAAACAAACAGCA	1280
	/protein_id="AAA67114.1"		1021	TTTGACTTAAATTTAGAAAAAGCTGAAACAGCAAACTAAACCACTTTTCACTGCCTGAC	1080
	/db_xref="GI:332296"		1281	TTTGACTTAAATTTAGAAAAAGCTGAAACAGCAAACTAAACCACTTTTCACTGCCTGAC	1340
	/translations="MSDGTSDPSGNVHSAVERAADPGSGGGSGGGVGVST		1081	ACAAGAACCTGCAGAAATTTTCTTTTCTGCTGCAATCTGTTAAAGTTTCCCATGCT	1140
	GSVDNTHRPLFDGQVEITALATRLVHLNPKSENYCRIRVHNTTDTSVKGNMAKDL		1341	ACAAGAACCTGCAGAAATTTTCTTTTCTGCTGCAATCTGTTAAAGTTTCCCATGCT	1400
	AHQEIWTPSLVDANAGVWLQSDMOYICNTMSQNLVSLDQSEIFNVVLTQTEQDL		1141	ATTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATACTGTTTATTTATTCATGACCA	1200
	GGQAIKIYNNDLTACMVAVDNSNNILPYTPAANSMETLGFYPWKPTIATSPYRYFCDV		1401	ATTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATACTGTTTATTTATTCATGACCA	1460
	RDLSVYENQEGTEVENVMTGPKMNSQFFTIENQOITLLRTGDBFATGTPYFDTS		1201	GCAGACAGGCAAACTCTATTTATTCAGCAAGCCATGACGCAAGAGTGGCAATGTTGGT	1320
	VKLTHWQNTNROLGQPLLSLTPPEADTDAGTLTAQGRHGTQMGVNMVSEAIRTPA		1461	GCAGACAGGCAAACTCTATTTATTCAGCAAGCCATGACGCAAGAGTGGCAATGTTGGT	1520
	QVGFCDPHNDPEASRAPPAAPKVPADITQGVNDKEANGSVSYGKHGHNWASHGPA		1261	TGCTATAATGCGACCAATCTGTTAACTTTTCCATTTTAAAGCTGTTACCAAGAACTTGAT	1380
	PERYTWDETSFGSRDTKGFIQSAPLVPPPLNGILTNANPLGTNDLHFNVSNSY		1521	TGCTATAATGCGACCAATCTGTTAACTTTTCCATTTTAAAGCTGTTACCAAGAACTTGAT	1580
	GLPTAFSPSPVTPQGOIWDKELDLEHKLPHLITAPFVCKNNAPGGQMLVRLGNLTDQ		1321	TGGGTAGAAAGAGCTGCTAACTTTTGACAGCAAGTAAACAGTCTTAAAGCTCATTTGCTCT	1380
			1581	TGGGTAGAAAGAGCTGCTAACTTTTGACAGCAAGTAAACAGTCTTAAAGCTCATTTGCTCT	1640
	Query Match	99.9%; Score 2017.4; DB 14; Length 5149;	1381	GFTCAAGCTATTGCGAATTGATCAAAAAGGAAAGGAGCAAAACAGATTGAACCAACCA	1440
Best Local Similarity	100.0%; Pred. No. 0;		1641	GFTCAAGCTATTGCGAATTGATCAAAAAGGAAAGGAGCAAAACAGATTGAACCAACCA	1700
Matches 2018; Conservative	0; Mismatches 1; Indels 0; Gaps 0;		1441	GTCATCATGACCAAAATGAGACATTAAGTGGTTCAGATAGGCTGCGAAGAAAGACCA	1500
QY	1	ATGCGTGGAAATGCTTACTGATGAGTGTGGAGCAACCACTGGTTAAAGGAAAAA	60	60	
DB	261	ATGCGTGGAAATGCTTACTGATGAGTGTGGAGCAACCACTGGTTAAAGGAAAAA	320	320	
QY	61	AGTAACGAGCAAGTCTCATTTGTTTAAATAAGTCTCAACTGGAATGGAATA	120	120	
DB	321	AGTAACGAGCAAGTCTCATTTGTTTAAATAAGTCTCAACTGGAATGGAATA	380	380	
QY	121	GATATCGGATGGAATAGTTACAAAAAGAGCTGCGAGGAGCAGAGTGAATCTTTACAA	180	180	
DB	381	GATATCGGATGGAATAGTTACAAAAAGAGCTGCGAGGAGCAGAGTGAATCTTTACAA	440	440	
QY	181	CGAGGAGCGGAACTACTTGGGACCAAGCGAGCAATGAATGGGAAACCAAGTGGAT	240	240	
DB	441	CGAGGAGCGGAACTACTTGGGACCAAGCGAGCAATGAATGGGAAACCAAGTGGAT	500	500	
QY	241	GAAATGACCAAAAGCAAGTATTCATTTTGTGTTTAAATAAGTCTTTATTTGAA	300	300	
DB	501	GAAATGACCAAAAGCAAGTATTCATTTTGTGTTTAAATAAGTCTTTATTTGAA	560	560	
QY	301	GTGCTTAAACCAAGAAATATTTCTGCTGATGTTAAATGTTTGTGCAACATGAATGG	360	360	
DB	561	GTGCTTAAACCAAGAAATATTTCTGCTGATGTTAAATGTTTGTGCAACATGAATGG	620	620	
QY	361	GGAAAGACCAAGCTGGCACTGCCATGTACTAATTTGAGGAAAGGACTTTAGTCAAGCT	420	420	
DB	621	GGAAAGACCAAGCTGGCACTGCCATGTACTAATTTGAGGAAAGGACTTTAGTCAAGCT	680	680	
QY	421	CAAGGGAATGTTGGAGAAAGCACTTAATTTTCTGAGGAGCAGATGGTTGTTAAACGCC	480	480	
DB	681	CAAGGGAATGTTGGAGAAAGCACTTAATTTTCTGAGGAGCAGATGGTTGTTAAACGCC	740	740	
QY	481	TGTAATGTGCAACTAACACCAAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT	540	540	
DB	741	TGTAATGTGCAACTAACACCAAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT	800	800	
QY	541	GAGTGGGTACTCTTACTTACTTAAAGCAATAAGCAAAACCAAAAGACTATACCAAGTGT	600	600	
DB	801	GAGTGGGTACTCTTACTTACTTAAAGCAATAAGCAAAACCAAAAGACTATACCAAGTGT	860	860	
QY	601	GTCTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAGAAATAAGCACTAGT	660	660	

QY 1741 TCACCAATTCAGCACCGAAAGTACGCCCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800
DB 2001 TCACCAATTCAGCACCGAAAGTACGCCCTCTCAGCCAGAACTATGCACTAACTCCACTT 2060
QY 1801 GCATCGATTCGAGGACCTGGCTTTAGAGCTTTGAGGACACCAAAATCTCTCTGTCG 1860
DB 2061 GCATCGATTCGAGGACCTGGCTTTAGAGCTTTGAGGACACCAAAATCTCTCTGTCG 2120
QY 1861 GGCACTCAGAAACCCAGAACACTGGGAAAGCTGGTTCCAAAGCTCCCAAGATGGTCAA 1920
DB 2121 GGCACTCAGAAACCCAGAACACTGGGAAAGCTGGTTCCAAAGCTCCCAAGATGGTCAA 2180
QY 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGTTGGGGAACCG 1980
DB 2181 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGTTGGGGAACCG 2240
QY 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACCTTGACTAA 2019
DB 2241 TTGAAGAAAGACTTCAGCGAGCGCTGAACCTTGACTAA 2279

RESULT 5
AX137739
LOCUS AX137739 2019 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 4 from Patent EPI077260.
ACCESSION AX137739
VERSION AX137739.1 GI:14273913
KEYWORDS
SOURCE Mice minute virus
ORGANISM Mice minute virus
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE 1
Nuesch, J. and Rommelaere, J.
AUTHORS
TITLE Parvovirus ns1 variants
JOURNAL Patent: EP 1077260-A 4 21-FEB-2001;
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts
(DE)

FEATURES
source Location/Qualifiers
1..2019
/organism="Mice minute virus"
/mol_type="unassigned DNA"
/db_xref="taxon:10794"
/note="parvovirus NS1 variant"
1..2019
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC39991.1"
/db_xref="GI:14273914"
/translation="MAGNAYSDVLGATNWLKESNQEVFSFVFKNENVQLNGKDIGW
NSYKLEODELSLORGAETTWQSEDMEWETTVDWMKKOVIIFDSLKKCLPEVL
NTKNIIPGDVNVFQHWGDKQCHVLIGKDFSOAQKWRQOLNVYWSRLVTA
CNQLPFAERIKUREIAEDNEWTLITRYHKYKDYTKCVLPFGNMIAYFLTKKIS
TSPRDGGVFLSDSSDKNFKLKEGERHLVSKLTDMPETVEITVTAQETKGR1
QTKVEAIKTLKELVHKRVTSPEWMMQPDYSIEMMAQPGGENLLKNTLEITLTL
ARTKTAFLDLEKAEYSKLFNSLPDTRICRIFAFGNWKVCHAI CCVLRQGGKR
NTVLFHPASTGKSI IQAIAOAVGNVCYNANVFPENDCNKLI WVEEAGNFGQ
QVNFKNICSGQIRIDQKGSQKLEPTPVIMTNENITVRI GCEEPBHTQPLRD
RLNLIHLTLPDGLVDKNFPMICAMLVKNGYSTWASYCAKMGKVPDWSNWE
PKVPTFINLIGARSFPTPKPSQNSVALTPLASDLLEALPEWSTNTPVAGTAE
TQNTGAGSKACQDQGLSPTWSIEEDLRACFGAEPLKKDFSEPLND"

ORIGIN
Query Match 99.88; Score 2015.8; DB 6; Length 2019;
Best Local Similarity 99.98; Pred. No. 0;
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGCTCGAAATGCTTACTCTGATGAAGTTTGGGACCAACCACTGGTTAAAGGAAAA 60
DB 1 ATGCTGGAATGCTTACTCTGATGAAGTTTGGGACCAACCACTGGTTAAAGGAAAA 60
QY 61 AGTAACGAGAGTGTCTCATTTGTTTTTAAAAATGTTCAACTGAATGAAAA 120

DB 61 AGTAACGAGAGTGTCTCATTTGTTTTTAAAAATGTTCAACTGAATGAAAA 120
QY 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCGAGGAGCAGAGCTGAAATCTTTACAA 180
DB 121 GATATCGGATGGAATAGTTTACAAAAAGAGCTGCGAGGAGCAGAGCTGAAATCTTTACAA 180
QY 181 CGAGGAGCGGAAAACCTACTTTGGGACCAAGCGAGGACATGGAATGGGAAACCCACAGTGGAT 240
DB 181 CGAGGAGCGGAAAACCTACTTTGGGACCAAGCGAGGACATGGAATGGGAAACCCACAGTGGAT 240
QY 241 GAATACCAAAAAGCAAGTATTCATTTTGGTTTATCTTCTTGGTTAAAAAATGTTTATTGAA 300
DB 241 GAATACCAAAAAGCAAGTATTCATTTTGGTTTATCTTCTTGGTTAAAAAATGTTTATTGAA 300
QY 301 GTCTTTAAACACAAAGAAATATATTTCTCGTGTGATTTAAATTTGGTTGCGAAACATGAATGG 360
DB 301 GTCTTTAAACACAAAGAAATATATTTCTCGTGTGATTTAAATTTGGTTGCGAAACATGAATGG 360
QY 361 GGAAGAGCAAGGCTGGCAGTGCCTCATGTA CTAATTTGGAGGAAAGCACTTTAGTCAAGCT 420
DB 361 GGAAGAGCAAGGCTGGCAGTGCCTCATGTA CTAATTTGGAGGAAAGCACTTTAGTCAAGCT 420
QY 421 CAAGGAAATGCTGGAGAGGCACTAAATGTTTACTGGAGCAGATGCTTGGTAA CAGCC 480
DB 421 CAAGGAAATGCTGGAGAGGCACTAAATGTTTACTGGAGCAGATGCTTGGTAA CAGCC 480
QY 481 TGTAAATGTGCAACTAACACACAGCTGAAAGAAATTAATACTAAGAGAAATAGCAGAGCAAT 540
DB 481 TGTAAATGTGCAACTAACACACAGCTGAAAGAAATTAATACTAAGAGAAATAGCAGAGCAAT 540
QY 541 GAGTGGTTTACTCTACTTATTAGCAATAAGCAAAACCAAAAGACATATACCAAGTGT 600
DB 541 GAGTGGTTTACTCTACTTATTAGCAATAAGCAAAACCAAAAGACATATACCAAGTGT 600
QY 601 GTTCTTTTGGAAAACATGATGCTTACTATTTTAACTAAAAAGAAATAGCACTACT 660
DB 601 GTTCTTTTGGAAAACATGATGCTTACTATTTTAACTAAAAAGAAATAGCACTACT 660
QY 661 CCACCAAGAGACGAGGCTATTTTCTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA 720
DB 661 CCACCAAGAGACGAGGCTATTTTCTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA 720
QY 721 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCGAGAAACG 780
DB 721 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCGAGAAACG 780
QY 781 GTTGAACACCAAGTAACTACCTGCGGAGGAACTAAGCGGCGAGATTCAAACTAAAAA 840
DB 781 GTTGAACACCAAGTAACTACCTGCGGAGGAACTAAGCGGCGAGATTCAAACTAAAAA 840
QY 841 GAAGTTTCTATTAAAACTACCTTAAAGAGCTGGTGCTATAAAAAGAGTAACTCACCAGAG 900
DB 841 GAAGTTTCTATTAAAACTACCTTAAAGAGCTGGTGCTATAAAAAGAGTAACTCACCAGAG 900
QY 901 GACTGGATGATGATGAGCAGCAGACAGTTACATTTGAAATGATGCTCAACACAGTGGAGAA 960
DB 901 GACTGGATGATGATGAGCAGCAGACAGTTACATTTGAAATGATGCTCAACACAGTGGAGAA 960
QY 961 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAACCAACAGCA 1020
DB 961 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAACCAACAGCA 1020
QY 1021 TTTGACTTAATTTTAAAAAAGCTGAAACAGCAAACTAACTAACTTTTCACTGCTCTGAC 1080
DB 1021 TTTGACTTAATTTTAAAAAAGCTGAAACAGCAAACTAACTAACTTTTCACTGCTCTGAC 1080
QY 1081 ACAAGAACCTGAGAAATTTTCTTTTATGCTGGAATATATGTTTAAAGTTTGGCCATGCT 1140
DB 1081 ACAAGAACCTGAGAAATTTTCTTTTATGCTGGAATATATGTTTAAAGTTTGGCCATGCT 1140
QY 1141 ATTTGCTGTGTTTTAAACAGACAGAGGCAAGAAATACTGTTTATTTCAATGAGCA 1200
DB 1141 ATTTGCTGTGTTTTAAACAGACAGAGGCAAGAAATACTGTTTATTTTCAATGAGCA 1200

Qy	1201	GCAGCAGCGGAAATCTATTTATTCACAGCCATAGCACAGCAGTTGGCAATGTTGGT	1260
Db	1201	GCAGCAGCGGAAATCTATTTATTCACAGCCATAGCACAGCAGTTGGCAATGTTGGT	1260
Qy	1261	TGCTATAATCAGCGCAATGTAACTTTCCATTTAATGACTGTACCAACAAGAACTTGATT	1320
Db	1261	TGCTATAATCAGCGCAATGTAACTTTCCATTTAATGACTGTACCAACAAGAACTTGATT	1320
Qy	1321	TGGGTAGAAGAAGCTGGTAACTTTGGACAGCAAGTAAACAGTTTAAAGCCATTGCTCT	1380
Db	1321	TGGGTAGAAGAAGCTGGTAACTTTGGACAGCAAGTAAACAGTTTAAAGCCATTGCTCT	1380
Qy	1381	GGTCAAGCTATTGCGATTGATCAAAAAGGAAAGGCGAGCAAAACAGATTGAACCAACCA	1440
Db	1381	GGTCAAACTATTGCGATTGATCAAAAAGGAAAGGCGAGCAAAACAGATTGAACCAACCA	1440
Qy	1441	GTCAATCATGACCAAAATGAGAACATTAACAGTGTGAGATAGCTGCGAAGAAAGACCA	1500
Db	1441	GTCAATCATGACCAAAATGAGAACATTAACAGTGTGAGATAGCTGCGAAGAAAGACCA	1500
Qy	1501	GAACACACTCAACCAATCAGAGACAGAAATGCTTAAACATTCATCTAACACATACCTTGCCT	1560
Db	1501	GAACACACTCAACCAATCAGAGACAGAAATGCTTAAACATTCATCTAACACATACCTTGCCT	1560
Qy	1561	GGTGACTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGCTTGGTTGGTAAAG	1620
Db	1561	GGTGACTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGCTTGGTTGGTAAAG	1620
Qy	1621	AATGTTTACCAATCTACATGGCAAGCTACTGTCTAAATGGGCGCAAGTTCTCTGATTGG	1680
Db	1621	AATGTTTACCAATCTACATGGCAAGCTACTGTCTAAATGGGCGCAAGTTCTCTGATTGG	1680
Qy	1681	TCAGAAAACCTGGGCGGAGCAAGGTCGCAACTCTATAAATTAATAGTTTGGCAGCG	1740
Db	1681	TCAGAAAACCTGGGCGGAGCAAGGTCGCAACTCTATAAATTAATAGTTTGGCAGCG	1740
Qy	1741	TCACCAATTCAGACACCGAAAAGTACGCTCTCAGCGAGAACTATGCACTAACTCCACTT	1800
Db	1741	TCACCAATTCAGACACCGAAAAGTACGCTCTCAGCGAGAACTATGCACTAACTCCACTT	1800
Qy	1801	GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACCAATCTCTCTGTTGG	1860
Db	1801	GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACCAATCTCTCTGTTGG	1860
Qy	1861	GGCACTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGGTCAA	1920
Db	1861	GGCACTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCCAAGATGGTCAA	1920
Qy	1921	CTGAGCCCAACTTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTCGGTGCGGAACG	1980
Db	1921	CTGAGCCCAACTTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTCGGTGCGGAACG	1980
Qy	1981	TTCAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA	2019
Db	1981	TTCAAGAAAGACTTCAGCGAGCGCTGAACTTGGACTAA	2019

RESULT 6	
AX137743	
LOCUS	AX137743
DEFINITION	Sequence 8 from Patent EP1077260.
ACCESSION	AX137743
VERSION	AX137743.1
KEYWORDS	GI:14273917
SOURCE	Mice minute virus
ORGANISM	Mice minute virus
REFERENCE	1
AUTHORS	Nuesch, J. and Rommelaere, J.
TITLE	Parvovirus ns1 variants
JOURNAL	Patent: EP 1077260-A 8 21-FEB-2001;
	Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts

FEATURES		(DB)	Location/Qualifiers	
source			1. .2019	
			/organism="Mice minute virus"	
			/mol_type="unassigned DNA"	
			/db_xref="taxon:10794"	
			/notes="Parvovirus ns1 variant"	
CDS			1. .2019	
			/note="unnamed protein product"	
			/codon_start=1	
			/protein_id="CAC39993.1"	
			/db_xref="GI:14273918"	
			/translation="MAGNAYSDEVLGATNWLKEKSQNEVPSFVKENVOLNKDIGW	
			NSTKELQDELKSLQRGAEITWDQSEMEWETTVDMEITKKQVIFDLSVKKCLCEVL	
			NTKNIPEFQVNMVFWHEGKQDQGHVLIIGKDFQAGQKWRRLNLYVWSRLVLA	
			CNVQLTPAERIKLREIAEDNEWLLTYKHQTKDYTKCVLFKNMIAIYVFLTKKITS	
			TSPRDGGVPLSSDSQKTNFLKGEGRHLVSKLYTDDMRPETVETVTTAQETKGRRI	
			QTKKEYIKITLKLKELVHRTVSPEDMMWQPDYSYIEMWQPGGENILLKNTLEICTLTL	
			ARTKTAFDLILLEKAETSKLTNPSLPDTRACRIPAPHGNVYVKVCHAICCVLNRQGGKR	
			NTVLFPHGPASTGKSIIAQALAAQAVGNVGTCAANVFPFNDCKNKLIVVEAGNFGQ	
			QVNFKAICSGQTIKIDQKSGSKQIEPTFVINTNENITVVRIGCEERPEHTOPIRD	
			RMLNIHLTHLPDGLVDKNWPMICAWLVKNQYQSTWASVCAKWKVDPDWNWAE	
			PKVPTPINLIGSARSPTTPKSTPLSONVALTPLASDLDELALPEWSTPNTPVAGTAE	
			TQNTGEGSKACQDQGLSPTWSIEEDLRACFCAEPLKXDFSEPLND"	
ORIGIN				
Query Match		99.8%;	Score 2015.8;	DB 6; Length 2019;
Best Local Similarity		99.9%;	Pred. No. 0;	
Matches 2017;		Conservative 0;	Mismatches 2;	Indels 0; Gaps 0;
Qy	1	ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA	60	
Db	1	ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA	60	
Qy	61	AGTAACACGAGAGCTGTTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGGAAAA	120	
Db	61	AGTAACACGAGAGCTGTTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGGAAAA	120	
Qy	121	GATATCGGATGGAAATAGTTACAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTACAA	180	
Db	121	GATATCGGATGGAAATAGTTACAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTACAA	180	
Qy	181	CGAGGAGCGGAAACTACTTGGGACCAAAAGGAGGACATGGAATGGGAAACACAGTGGAT	240	
Db	181	CGAGGAGCGGAAACTACTTGGGACCAAAAGGAGGACATGGAATGGGAAACACAGTGGAT	240	
Qy	241	GAAATGACCAAAAGCAAGTATTCTTTGTTGTTTAAATAAATAAATGTTTAAATTTGAA	300	
Db	241	GAAATGACCAAAAGCAAGTATTCTTTGTTGTTTAAATAAATAAATGTTTAAATTTGAA	300	
Qy	301	GTGCTTAAACACAAAGAAATATATTTCTGCTGATGTTAAATTTGTTGCAACATGAATGG	360	
Db	301	GTGCTTAAACACAAAGAAATATATTTCTGCTGATGTTAAATTTGTTGCAACATGAATGG	360	
Qy	361	GGAAAAAGCAAGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGCACTTTAGTCAAGCT	420	
Db	361	GGAAAAAGCAAGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGCACTTTAGTCAAGCT	420	
Qy	421	CAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGGTTGTTAAACAGCC	480	
Db	421	CAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGGTTGTTAAACAGCC	480	
Qy	481	TGTAATGTGCAACTAACACCAAGCTGAAAGAAATTAAGAGAAATAGCAGAGACAAT	540	
Db	481	TGTAATGTGCAACTAACACCAAGCTGAAAGAAATTAAGAGAAATAGCAGAGACAAT	540	
Qy	541	GAGTGGTTTACTTCTTACTTATAGCATAAGCAAAACCAAAAGACATATACCAAGTGT	600	
Db	541	GAGTGGTTTACTTCTTACTTATAGCATAAGCAAAACCAAAAGACATATACCAAGTGT	600	
Qy	601	GTTCTTTTGGAAACATGATTCGTACTATTTTAACTAAAAAGAAATAAGCACTAGT	660	
Db	601	GTTCTTTTGGAAACATGATTCGTACTATTTTAACTAAAAAGAAATAAGCACTAGT	660	

661 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA 720
Db
661 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA 720
Qy
721 AAAGAAGGCGAGCCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAAG 780
Db
721 AAAGAAGGCGAGCCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAAG 780
Qy
781 GTTGAACACACAGTAAACCTAGCGAGGAACTTAAGCGGCGCAGAACTCAAACTAAAAA 840
Db
781 GTTGAACACACAGTAAACCTAGCGAGGAACTTAAGCGGCGCAGAACTCAAACTAAAAA 840
Qy
841 GAAAGTTTCTATTAAAACTACACTTAAAGAGCTGGTGATATAAGAGCTTAACCTCACCAGAG 900
Db
841 GAAAGTTTCTATTAAAACTACACTTAAAGAGCTGGTGATATAAGAGCTTAACCTCACCAGAG 900
Qy
901 GACTGGATGATGATGACGACGACAGTTACATTGAAATGATGGCTCAACAGGTGAGAA 960
Db
901 GACTGGATGATGATGACGACGACAGTTACATTGAAATGATGGCTCAACAGGTGAGAA 960
Qy
961 AACCTGCTGAAAAATAGCTAGAGATTGTACACTAACTAGCCAGAACCAAAACAGCA 1020
Db
961 AACCTGCTGAAAAATAGCTAGAGATTGTACACTAACTAGCCAGAACCAAAACAGCA 1020
Qy
1021 TTTGACTTAATTTAGAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCTGAC 1080
Db
1021 TTTGACTTAATTTAGAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCTGAC 1080
Qy
1081 ACAAGAACTGCAAGATTTTGTCTTTTCATGCTGGAATATGTTAAAGTTTGCCTATGCT 1140
Db
1081 ACAAGAACTGCAAGATTTTGTCTTTTCATGCTGGAATATGTTAAAGTTTGCCTATGCT 1140
Qy
1141 ATTTGCTGTTTTTAAACACAGAGGAGCAAGAAAGAACTACTGTTTTTATTCATGACCA 1200
Db
1141 ATTTGCTGTTTTTAAACACAGAGGAGCAAGAAAGAACTACTGTTTTTATTCATGACCA 1200
Qy
1201 GCCAGACAGGCAAACTTATTATGCAAGCCATAGCAACAGCAGTTGGCAATTTGGT 1260
Db
1201 GCCAGACAGGCAAACTTATTATGCAAGCCATAGCAACAGCAGTTGGCAATTTGGT 1260
Qy
1261 TGCTATAATGACGCAATGTAATCTTTTCAATTAATGATGATGACCAAGAACTGAT 1320
Db
1261 TGCTATAATGACGCAATGTAATCTTTTCAATTAATGATGATGACCAAGAACTGAT 1320
Qy
1321 TGGGTAGAGAAAGCTGTAATCTTTGACAGCAAGTAAACAGTTTAAAGCAATTTGCTCT 1380
Db
1321 TGGGTAGAGAAAGCTGTAATCTTTGACAGCAAGTAAACAGTTTAAAGCAATTTGCTCT 1380
Qy
1381 GGTCAAGCTATTGCGATTGATCAAAAAGGAAAGGAGGAGCAAAAGATTTGAACCAACCA 1440
Db
1381 GGTCAAGCTATTGCGATTGATCAAAAAGGAAAGGAGGAGCAAAAGATTTGAACCAACCA 1440
Qy
1441 GTCATATGACCAAAATGAGAACATTAAGTGGTGAATAGGCTGCGAAGAAAGACCA 1500
Db
1441 GTCATATGACCAAAATGAGAACATTAAGTGGTGAATAGGCTGCGAAGAAAGACCA 1500
Qy
1501 GAACACACTCAACCAATCAGAGACAGATGCTTAAACATTTCAATCTAACATACATCTGCT 1560
Db
1501 GAACACACTCAACCAATCAGAGACAGATGCTTAAACATTTCAATCTAACATACATCTGCT 1560
Qy
1561 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGTGCTTGGTTGGTAAAG 1620
Db
1561 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGTGCTTGGTTGGTAAAG 1620
Qy
1621 AATGGTTACCAATCTACATGCGAGCTACTGTGCTAAATGGGCAAGTTCTGATTGG 1680
Db
1621 AATGGTTACCAATCTACATGCGAGCTACTGTGCTAAATGGGCAAGTTCTGATTGG 1680
Qy
1681 TCAGAAACTGGGCGGAGCAAGGTGCTTAAATTTACTTAAATTTACTAGGTTGGGACGC 1740
Db
1681 TCAGAAACTGGGCGGAGCAAGGTGCTTAAATTTACTTAAATTTACTAGGTTGGGACGC 1740

1741 TCACCATTCACACACCGGAAAAAGTAGCGCTCTCAGCCAGAACTATGCACTAACTCAGCTT 1800
Db
1741 TCACCATTCACACACCGGAAAAAGTAGCGCTCTCAGCCAGAACTATGCACTAACTCAGCTT 1800
Qy
1801 GCATCGGATCTCGAGGACCTGCTTTAGAGCCTTGGAGCACACCAATACTCCTGTTGG 1860
Db
1801 GCATCGGATCTCGAGGACCTGCTTTAGAGCCTTGGAGCACACCAATACTCCTGTTGG 1860
Qy
1861 GGCACCTGCAAGAAACCCAGAACTGCGGGAAGCTGGTTCCAAAGCCTGCCAAGATGCTCAA 1920
Db
1861 GGCACCTGCAAGAAACCCAGAACTGCGGGAAGCTGGTTCCAAAGCCTGCCAAGATGCTCAA 1920
Qy
1921 CTGAGCCCAACTGCTGCTAGAGATCGAGGAGGATTTGAGAGCGTCTCGGTCGGGAACCG 1980
Db
1921 CTGAGCCCAACTGCTGCTAGAGATCGAGGAGGATTTGAGAGCGTCTCGGTCGGGAACCG 1980
Qy
1981 TTGAAGAAAGACTTCAGGAGCGCGCTGAACTTTGGACTAA 2019
Db
1981 TTGAAGAAAGACTTCAGGAGCGCGCTGAACTTTGGACTAA 2019

RESULT 7
AX137747
LOCUS AX137747 2019 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 12 from Patent EP1077260.
ACCESSION AX137747
VERSION AX137747.1 GI:14273921
KEYWORDS
SOURCE
ORGANISM
Mice minute virus
Mice minute virus
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE
1 Nueesch, J. and Rommelaere, J.
Parvovirus ns1 variants
Patent: EP 1077260-A 12 21-FEB-2001;
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts
(DE)
FEATURES
Location/Qualifiers
1..2019
/organism="Mice minute virus"
/mol_type="unassigned DNA"
/db_xref="taxon:10794"
/note="Parvovirus NS1 variant"
1..2019
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC39995.1"
/db_xref="GI:14273922"
/translation="MAGNAYSDVLGATNWLKESNOEVFSVFKNENVLNGKDIGM
NSYKXELQDELSLQAGATTTWQSEDHNEWETTVDEMTKKQVIFDSLVKKCLFEVL
NTKNTPPGDVNFWQHGKQDQWHCVLIGKDFGSAQKQKWRRLQNLVYNSRWLVTA
CNVQLTPABRIKLREIAEDNEWVTLTYKHQTKKDYTKCVLFGNMIAYYFLTKKIS
TSPRPDGGVFLSSDGSWTKFLKEGRLHVLKLYTDMRPETVETVTTVAQETKRGRI
OTPKVSIKTLTKELVHKVTSPEMMQPDYSYEMMAQPGGLENLKNLTLEICTLTL
ARTKAPDLLEKETSRLTNPSLPDTRTCRIFAHGWYVKVCHALCVLNRQGGKR
NAVLPFGPATGKSIITAIQAAGVNGVCYNANVNFNDNCNKLIVWEAGNFGQ
QWPKAICSGQTRIDQKGSQKIEPTVMTNTNENITVVRIGEEPEHTQPIRD
RMLNHLHTLPGDFGLVDKNWPMICAWLVKNGYOSTWASYCAKQKVPDWSNWA
PKVPTPINLLGARSPTTPKSTPLSQNYALTPLASDLEDLALPEWSTNTPVAGTAE
TQNTGBAGSKACQDGLSPTWSEIEDLRACFGAEPLKKDFSEPLND"

CDS
Query Match 99.8%; Score 2015.8; DB 6; Length 2019;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGTTTAAAGGAAAA 60
Db 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGTTTAAAGGAAAA 60
Qy 61 AGTAACACAGGAGTGTCTCAITTTTAAAAATGAAAAATGTTTCAACTGAATGAAAA 120
Db 61 AGTAACACAGGAGTGTCTCAITTTTAAAAATGAAAAATGTTTCAACTGAATGAAAA 120

121 GATATCGGATGGATAGTTACAAAAAGAGCTGCGAGGAGCGAGCTGAAATCTTTACAA 180
121 GATATCGGATGGATAGTTACAAAAAGAGCTGCGAGGAGCGAGCTGAAATCTTTACAA 180
181 CGAGGAGCGGAAACTACTTGGGACCAAGCGAGGACATGAATGGGAAACACACAGTGGAT 240
181 CGAGGAGCGGAAACTACTTGGGACCAAGCGAGGACATGAATGGGAAACACACAGTGGAT 240
241 GAAATGACCAAAAGCAAGTATTCATTTTGAATCTTTGGTTTAAATAATGTTTATTGGAA 300
241 GAAATGACCAAAAGCAAGTATTCATTTTGAATCTTTGGTTTAAATAATGTTTATTGGAA 300
301 GTGCTTAAACCAAGAAATATTTCTGCTGGTGAATTTAAATTTGTTTAAATAATGTTTATTGG 360
301 GTGCTTAAACCAAGAAATATTTCTGCTGGTGAATTTAAATTTGTTTAAATAATGTTTATTGG 360
361 GGAAGACCAAGGCTGCGACTGCGATGATTAATTTGGAGGAAAGGACTTTTGTCAAGCT 420
361 GGAAGACCAAGGCTGCGACTGCGATGATTAATTTGGAGGAAAGGACTTTTGTCAAGCT 420
421 CAAGGGAATTTGGTGGAGAGGCAAACTAAATGTTTACTGGAGCAGATGTTGGTAACAGCC 480
421 CAAGGGAATTTGGTGGAGAGGCAAACTAAATGTTTACTGGAGCAGATGTTGGTAACAGCC 480
481 TGTAATGTGCAACTTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT 540
481 TGTAATGTGCAACTTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGCAAT 540
541 GAGTGGGTTACTTACTTACTTATAGCATAAGCAACCAAAAGAAAGTACTTACCAAGTGT 600
541 GAGTGGGTTACTTACTTACTTATAGCATAAGCAACCAAAAGAAAGTACTTACCAAGTGT 600
601 GTTCTTTTGGAAACATGATGTTGCTTACTATTTTTTAACTAAAAAGAAATAAGCACTAGT 660
601 GTTCTTTTGGAAACATGATGTTGCTTACTATTTTTTAACTAAAAAGAAATAAGCACTAGT 660
661 CCAACGAGAGCGGAGCTATTTTCTTAGCAGTGAATCTGGCTGGGAAACCTAACTTTTAA 720
661 CCAACGAGAGCGGAGCTATTTTCTTAGCAGTGAATCTGGCTGGGAAACCTAACTTTTAA 720
721 AAAGAGGCGGAGCGCCATCTAGTGAGCAAACTACTACATGATGATGATGATGATGATGATG 780
721 AAAGAGGCGGAGCGCCATCTAGTGAGCAAACTACTACATGATGATGATGATGATGATGATG 780
781 GTTGAAACCAAGTAACTGCGGAGGAACTAAAGCGCGGAGAAATCAAACTAAAGAA 840
781 GTTGAAACCAAGTAACTGCGGAGGAACTAAAGCGCGGAGAAATCAAACTAAAGAA 840
841 GAAATTTCTAATTAATACTACATTAAGAGCTGGTGATTAAGAGAGTAACTCACCAGAG 900
841 GAAATTTCTAATTAATACTACATTAAGAGCTGGTGATTAAGAGAGTAACTCACCAGAG 900
901 GACTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
901 GACTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
961 AACTGCTGAAAAATAGCTAGAGATTTGTACACTAATCTAGCGAGAACCAAAACAGCA 1020
961 AACTGCTGAAAAATAGCTAGAGATTTGTACACTAATCTAGCGAGAACCAAAACAGCA 1020
1021 TTTGACTTAATTTTAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCTGAC 1080
1021 TTTGACTTAATTTTAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCTGAC 1080
1081 ACAAGAACCTGCAAGATTTTGTGCTTTTATGCTGGAATATGTTTAAAGTTTGGCTGCT 1140
1081 ACAAGAACCTGCAAGATTTTGTGCTTTTATGCTGGAATATGTTTAAAGTTTGGCTGCT 1140
1141 ATTGCTGCTGTTTAAACACAGAGGAGGCAAAAGAAATCTGTTTATTTATGACCA 1200
1141 ATTGCTGCTGTTTAAACACAGAGGAGGCAAAAGAAATCTGTTTATTTATGACCA 1200

QY 1201 GCCAGCACAGGCAAAATCTATTATTGCACAGCCATAGCAACAAGCTTGGCAATGTGGT 1260
Db 1201 GCCAGCACAGGCAAAATCTATTATTGCACAGCCATAGCAACAAGCTTGGCAATGTGGT 1260
QY 1261 TGCTATAATGACAGCAATGTAAACTTTTCACTTTTAAAGTGTACCAACAAGAACTTCATT 1320
Db 1261 TGCTATAATGACAGCAATGTAAACTTTTCACTTTTAAAGTGTACCAACAAGAACTTCATT 1320
QY 1321 TGGGTAGAGAAAGCTGGTAACTTTTGGACAGCAAGTAAACCCAGTTTAAAGCAATTTGCTCT 1380
Db 1321 TGGGTAGAGAAAGCTGGTAACTTTTGGACAGCAAGTAAACCCAGTTTAAAGCAATTTGCTCT 1380
QY 1381 GGTCAAGCTATTTCGCATTTGATCAAAAGGAAAGGAGGAGCAACAGATTGAAACCAACCA 1440
Db 1381 GGTCAAGCTATTTCGCATTTGATCAAAAGGAAAGGAGGAGCAACAGATTGAAACCAACCA 1440
QY 1441 GTCATCATGACCAAAATGAGAACATTTACAGTGGTCAAGATAGGCTGCGAAGAAAGACCA 1500
Db 1441 GTCATCATGACCAAAATGAGAACATTTACAGTGGTCAAGATAGGCTGCGAAGAAAGACCA 1500
QY 1501 GAAACACACTCAACCAATTCAGAGACAGAAATGCTTTAAACATTCATCTAACACATACCTTGCCT 1560
Db 1501 GAAACACACTCAACCAATTCAGAGACAGAAATGCTTTAAACATTCATCTAACACATACCTTGCCT 1560
QY 1561 GGTGACTTTGGTGGTGGTTCACAAAATGAATGSCCCATGATTTTGTGCTTGGTAAAG 1620
Db 1561 GGTGACTTTGGTGGTGGTTCACAAAATGAATGSCCCATGATTTTGTGCTTGGTAAAG 1620
QY 1621 AATGTTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTCTGATTGG 1680
Db 1621 AATGTTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTCTGATTGG 1680
QY 1681 TCAGAAAACTGGGCGGAGCCAAAGTGGCCAACTCTATATAATTTACTAGTTCGGCACGC 1740
Db 1681 TCAGAAAACTGGGCGGAGCCAAAGTGGCCAACTCTATATAATTTACTAGTTCGGCACGC 1740
QY 1741 TCACCAATTCAGACACGAAAGTACGCTCTCAGCAGAACTATGCACTAACTCCACTT 1800
Db 1741 TCACCAATTCAGACACGAAAGTACGCTCTCAGCAGAACTATGCACTAACTCCACTT 1800
QY 1801 GCATCGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGACACACCAATATCTCTCTGTTGCG 1860
Db 1801 GCATCGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGACACACCAATATCTCTCTGTTGCG 1860
QY 1861 GGCACCTGCAGAAACCCAGAACACTGGGGAGCTGGTTCCAAAGCTGCCAAGATGTGCAA 1920
Db 1861 GGCACCTGCAGAAACCCAGAACACTGGGGAGCTGGTTCCAAAGCTGCCAAGATGTGCAA 1920
QY 1921 CTGAGCCCAACTTTGGTTCAGAGATCGAGGAGATTTGAGAGCGTCTTCGGTGGGAACCG 1980
Db 1921 CTGAGCCCAACTTTGGTTCAGAGATCGAGGAGATTTGAGAGCGTCTTCGGTGGGAACCG 1980
QY 1981 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019
Db 1981 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019

RESULT 8
MWICG
LOCUS
DEFINITION
Minute virus of mice (MVM(i)), a lymphotropic variant of MVM,
complete genome.
ACCESSION
M12032
VERSION
M12032.1 GI:332289
KEYWORDS
alternative splicing; capsid protein; complete genome;
nonstructural protein.
SOURCE
Mice minute virus
ORGANISM
Mice minute virus
REFERENCE
1 (bases 1 to 5085)
Astell,C.R., Gardner,E.M. and Tattersall,P.
AUTHORS
DNA sequence of the lymphotropic variant of minute virus of mice,
MVM(i), and comparison with the DNA sequence of the fibrotropic

ORIGIN 1085 bp upstream of EcoRI site.									
Query Match 94.4%; Score 1905.4; DB 14; Length 5085;									
Best Local Similarity 96.5%; Pred. No. 0;									
Matches 1948; Conservative 0; Mismatches 71; Indels 0; Gaps 0;									
QY	1	ATGCTGGAAATCCTTACTCTGATGAGTTTTCGGAGCAACCACTGGTTAAAGGAA	60						
DB	262	ATGGCTGGAAATCCTTACTCTGATGAGTTTTCGGAAACCAACCACTGGTTAAAGGAA	321						
QY	61	AGTAACCAAGAGTGTCTCATTTGTTTAAATAAGAAATGTTCAACTGAATGCA	120						
DB	322	AGTAACCAAGAGTGTCTCATTTGTTTAAACTGAGGATGTTCAACTAAATG	381						
QY	121	GATATCGATGGAATAGTTACAAAAGAGCTCGAGGAGCAGAGTGAATCTTTAC	180						
DB	382	GATATCGATGGAATTAACAAAAGAGCTCGAGGAGCAGAGTGAATCTTTAC	441						
QY	181	CGAGGAGCGAAATCTTGGACCAAGCGAGGACATGGAATGGGAAACCAAGTGG	240						
DB	442	CGAGGAGCGAAATCTTGGACCAAGCGAGGACATGGAATGGGAAATCTA	501						
QY	241	GAAATGACCAAAAAGCAAGTATTCATTTTGTGTTTAAATAAGTGTATTTG	300						
DB	502	GAAATGACCAAAAAGCAAGTATTCATTTATGATCTCTTATGTTAAATAAG	561						
QY	301	GTGCTTAACCAAAAGATATTTCTGCTGATGTTAAATGGTTGTGCAACATGA	360						
DB	562	GTGCTTAGCACAAAATAATAGCTCTGCTGATGTTACTTTGTTGACGATGA	621						
QY	361	GGAAAGACCAAGCTGGCATGCTGATTAATTTGGAGGAAGAGCTTTAGTCA	420						
DB	622	GGAAAGACCAAGCTGGCATGCTGATTAATTTGGAGGAAGAGCTTTAGTCA	681						
QY	421	CAAGGGAATGTTGGAGAGCACTAAATGTTTACTTGGAGCAGATGTTGTTA	480						
DB	682	CAAGGGAATGTTGGAGAGCACTAAATGTTTACTTGGAGCAGATGTTGTTA	741						
QY	481	TGTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAG	540						
DB	742	TGTAATGTGCACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAG	801						
QY	541	GAGTGGTTACTCTTAATTAGCATAGCAAAACCAAAAGACTATATCAAGT	600						
DB	802	GAGTGGTTACTCTTAATTAGCATAGCAAAACCAAAAGACTATATCAAGT	861						
QY	601	GTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAGAAATTAAGCA	660						
DB	862	GTCTTTTGGAAATGATGCTTACTATTTTAACTAAAGAAATTAAGCA	921						
QY	661	CCACCAAGAGACGAGGCTATTTCTTAGCAGTGACTCTGGCTGGAAACTA	720						
DB	922	CCGCAAGGACGAGGCTATTTCTTAAGCAGTGACTCTGGCTGGAAACTA	981						
QY	721	AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACATGATGACATGCG	780						
DB	982	AAAGAGGCGAAACCCATCTAGTGAGCAAACTATACATGATGACATGCG	1041						
QY	781	GTTGAAACCAAGTAAACCTGCGCAGCAAACTAAGCGCGCAGATTTCA	840						
DB	1042	GTTGAAACCAAGTAAACCTGCGCAGCAAACTAAGCGCGCAGATTTCA	1101						
QY	841	GAAGTTTCTATTAACCTTAAAGAGCTGGTGATTAAGAGAGTAACCTC	900						
DB	1102	GAGTTTCTATTAACCACTTAAGAGCTAGTGATTAAGAGTAACCTC	1161						
QY	901	GACTGGATGATGACGAGCAGTATTAATGAAATGATGGCTCAACCAAGT	960						
DB	1162	GACTGGATGATGACGAGCAGTATTAATGAAATGATGGCTCAACCAAGT	1221						
QY	961	AACCTGCTGAAAATAACGCTAGAGATTTGTACATACTCTAGCCAGAA	1020						
DB	1222	AACCTGCTGAAAATAACGCTAGAGATTTGTACATACTCTAGCCAGAA	1281						

RESULT 9
MOU34253
LOCUS

MOU34253 4764 bp DNA linear VRL 21-AUG-1996

```
DEFINITION Mouse parvovirus 1b DNA.
ACCESSION U34253
VERSION U34253.1 GI:1464793
KEYWORDS
SOURCE Mouse parvovirus 1b
ORGANISM Mouse parvovirus 1b
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE 1 (bases 1 to 4764)
AUTHORS Besselsen, D.G., Pintel, D.J., Purdy, G.A., Besch-Williford, C.L., Franklin, C.L., Hook, R.R. Jr. and Riley, L.K.
TITLE Molecular characterization of newly recognized rodent parvoviruses
JOURNAL J. Gen. Virol. 77 (Pt 5), 895-911 (1996)
MEDLINE 96201434
PUBMED 8609486
REFERENCE 2 (bases 1 to 4764)
AUTHORS Besselsen, D.G.
TITLE Direct Submission
JOURNAL Submitted (17-AUG-1995) David G. Besselsen, Department of Veterinary Pathology, University of Missouri-Columbia, W213 Veterinary Medicine Building, Columbia, MO 65211, USA
FEATURES
source
1..4764
/organism="Mouse parvovirus 1b"
/mol_type="genomic DNA"
/specific_host="Mus musculus"
/db_xref="taxon:42841"
ORIGIN
Query Match 94.2%; Score 1902.2; DB 14; Length 4764;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 1946; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
QY 1 ATGCTGGAATGCTTACTCTGATGAAGTTTGGGACCAACCACTGGTTAAAGGAAAA 60
DB 122 ATGCTGGAATGCTTACTCTGATGAAGTTTGGGACCAACCACTGGTTAAAGGAAAA 181
QY 61 AGTAACGAGGAAGTTCTCATTTGTTTTTAAATAAGTAATGTTCAACTGGAATGGAATA 120
DB 182 AGTAACGAGGAAGTTCTCATTTGTTTTTAAATAAGTAATGTTCAACTGGAATGGAATA 241
QY 121 GATATCGGATGGATAGTATCAAAAAGAGCTGCGAGGACGAGCTGGAATCTTTACAA 180
DB 242 GATATCGGATGGATAGTATCAAAAAGAGCTGCGAGGACGAGCTGGAATCTTTACAA 301
QY 181 CGAGGACGGAATACTCTGGACCAAGCGAGACATGGAATGGGAAAAACACAGTGGAT 240
DB 302 CGAGGACGGAATACTCTGGACCAAGCGAGACATGGAATGGGAAATCTACAGTGGAT 361
QY 241 GAAATGCCAAAAGCAAGTATTCATTTGTTTGGTTTAAATAAGTATTTATTTGAA 300
DB 362 GAAATGCCAAAAGCAAGTATTCATTTGTTTGGTTTAAATAAGTATTTATTTGAA 421
QY 301 GTGCTTAAACAAGATATATTCCTGGTGATGTTAAATTTGGTTTGGCAACATGATGG 360
DB 422 GTGCTTAAACAAGATATATTCCTGGTGATGTTAAATTTGGTTTGGCAACATGATGG 481
QY 361 GGAAGAACAAGCTGCGCACTGCGATGATTAATTCGAGGAAGAGCTTTAGTCAAGCT 420
DB 482 GGGGAAGACCAAGCTGCGCACTGCGATGATTAATTCGAGGAAGAGCTTTAGTCAAGCT 541
QY 421 CAAGGGAAATGGTGGAGAGGCAACTAAATTTTACTTGGAGCAGATGGTTGTAACAGCC 480
DB 542 CAAGGGAAATGGTGGAGAGGCAACTAAATTTTACTTGGAGCAGATGGTTGTAACAGCC 601
QY 481 TGTAAATGCAACTAAACACCAAGCTGAAGATTAATACTAAGAGAAATAGCAGAGCAAT 540
DB 602 TGTAAATGCAACTAAACACCAAGCTGAAGATTAATACTAAGAGAAATAGCAGAGCAAT 661
QY 541 GAGTGGGTACTCTACTTACTTAAAGCATAAGCAACCAAAAAGACTATACCAAGTGT 600
DB 662 GAGTGGGTACTCTACTTACTTAAAGCATAAGCAACCAAAAAGACTATACCAAGTGT 721
QY 601 GTTCTTTTGGAAACATGATGCTTACTATTTTTTAACTAAAAAGAAATAAGCACTAGT 660
```

Db 1802 ACAGAAACTGGCGGAGCGAAGGTCCGACTCTCTATATAATTCACTAGGTTCCGACGC 1861
Qy 1741 TCACCAATTCAGACACCGAAAAGTACGCTCTCAGCAGAACATATGCACTAATCCACTT 1800
Db 1862 TCACCAATTCAGACACCGAAAAGTACGCTCTCAGCAGAACATATGCACTAATCCACTT 1921
Qy 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGAGAGCACACCAATATCTCCTGTTGCG 1860
Db 1922 GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGAGAGCACACCAATATCTCCTGTTGCG 1981
Qy 1861 GGCACTCAGAAAACCCAGAACACTGGGGAAGCTGTTTCCAAAGCCTGCCAAGATGGTCAA 1920
Db 1982 GGCACTCAGAAAACCCAGAACACTGGGGAAGCTGTTTCCAAAGCCTGCCAAGATGGTCAA 2041
Qy 1921 CTGAGCCCAACTGGTTCAGAGATCGAGAGGATTTGAGAGGCTGTTCCGTTGCGGAACCG 1980
Db 2042 CTGAGCCCAACTGGTTCAGAGATCGAGAGGATTTGAGAGGCTGTTCCGTTGCGGAACCG 2101
Qy 1981 TTGAAGAAGACTTCAGGAGCGCTCAACTTGGACTAA 2019
Db 2102 TTGAAGAAGACTTCAGGAGCGCTCAACTTGGACTAA 2140

RESULT 10

PAVMI PAVMI 5087 bp DNA linear VRL 01-JUL-1999
LOCUS Mouse parvovirus minute virus immunosuppressive variant genome (= PAVMI).
DEFINITION X02481.1 GI:60918
ACCESSION X02481
VERSION X02481
KEYWORDS coat protein; genome; origin of replication; overlapping genes; terminal repeat; unidentified reading frame.
SOURCE Mouse minute virus
ORGANISM Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE 1 (bases 1 to 5087)
AUTHORS Sahli,R., McMaster,G.K. and Hirt,B.
TITLE DNA sequence comparison between two tissue-specific variants of the autonomous parvovirus, minute virus of mice
JOURNAL Nucleic Acids Res. 13 (10), 3617-3633 (1985)
MEDLINE 85242059
PUBMED 3855242
COMMENT For the fibroblast-specific strain (MVMp) sequence see <PAVMW2>. The genomes of MVMp and MVMi (immunosuppressive variant) have more than 96% of their sequence in common.
Data kindly reviewed (18-JUL-1986) by G. McMaster.

FEATURES

source 1..5087
/organism="Mouse minute virus"
/mol_type="genomic DNA"
/db_xref="taxon:10794"
misc_feature 1..116
/notes="terminal hairpin region"
promoter 176..183
/note="TATA box"
gene 2405..4558
/gene="VP1"
CDS <2405..4558
/gene="vp1"
/function="coat protein"
/codon_start=1
/product="vp1 protein"
/protein_id="CAB46507.1"
/db_xref="GI:5419928"
/db_xref="GOA:P07302"
/db_xref="Swiss-Prot:P07302"
/translation="VPPGYLPGNSLDGSEPTNPSDAAKEHDEAYDQYIKSGRNP
YLYFSADQRFIDQTKADMGKGVHYFFRTKRAFAPKLADSPEFTSGVSRAGRIT
RPAYTIPINQARAKLTSSAAQSSQSDGTSPQDGGNVHSAARVERAADPGGS
GGGGGGGGVSTGYDNOYHRYFLGDGWEITATLATRLVHLNMPKSENYCRIRVHN
TTDTSVKGNWAKDDAHEQIWPMSLDANAWGLQPSDWQYICNTMSQLNLSLDOE
IPNVLTVEQDSGGQAIKYNNDLTACMMVAVDNSNLLPYTPAANSMTLGFYFWK
FTIASPYRYFYFCVDRDLSVTYENQBGITIEHNVNMGTPKGMNSQFFTIENQIITLLRTG

DEPATCTVYEDTNPKVLTHWTOTNROLGPPPLSTPREADTDTAGTLLTAQSRBQATOM
EUNVSEAIITRPAQVGFQCPQHNDFEASRAGPFAAPKVADVTQGVDRENGSRVSY
GKHGHNWAHGPAPRYTWDETFSGRDRDGFQSAFLVPPPLUNGILLTANPITG
TKNDIHFSNVNSYGLTAPSHSPYPOQIWDKELDLHKPLHITAPFVCKNNAP
GQMLVRLGNLTDQYDPNGATLSRIYVTGTFWFKGLTMRKLRANTTNPNVYQVSVE
DNGNSYMSVTWKLPATATGNMQSVPLITRPVARNY"
2792..4558
/gene="VP2"
CDS <2792..4558
/gene="VP2"
/function="coat protein"
/codon_start=1
/product="VP2 protein"
/protein_id="CAB46508.1"
/db_xref="GI:5419929"
/db_xref="GOA:P07302"
/db_xref="Swiss-Prot:P07302"
/translation="TMSDGTSPDGGNVHSAARVERAADPGSGGGGGGGVGVSY
TGSYDNOYHRYFLGDGWEITATLATRLVHLNMPKSENYCRIRVHNTDTSVKGNWAKD
DAHEQIWPMSLDANAWGLQPSDWQYICNTMSQLNLSLDOEIPNVVLKTVTEOD
SGQAIKYNNDLTACMMVAVDNSNLLPYTPAANSMTLGFYFWKPLHITAPFVCKNNAP
DRLSVTYENQBGITIEHNVNMGTPKGMNSQFFTIENQIITLLRTGDEFAITGYFDTN
PVKLTHTWOTNROLGPPPLSTPREADTDTAGTLLTAQSRBQATOMVNVWSEAIRTP
AOVGFCOPHNDFEASRAGPFAAPKVADVTQGVDRENGSRVSYGKHGHNWAHGP
APERITYWDETNFSGRDRDGFQSAFLVPPPLUNGILLTANPITGKNDIHFSNVNS
YGLTAPSHSPYPOQIWDKELDLHKPLHITAPFVCKNNAPGQMLVRLGNLTD
QYDPNGATLSRIYVTGTFWFKGLTMRKLRANTTNPNVYQVSVEDNGNSYMSVTWK
LPATATGNMQSVPLITRPVARNY"
4603..4608
misc_feature
/note="polyadenylation signal"
4821..4826
misc_feature
/note="polyadenylation signal"
4878..5087
misc_feature
/note="terminal hairpin region"

ORIGIN

Query Match 94.2%; Score 1902.2; DB 14; Length 5087;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 1946; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
Qy 1 ATGGCTGGAAATGCTTACTCTGATGAAGTGTTCGGGAGCAACCACTGGTTTAAAGGAAAA 60
Db 262 ATGGCTGGAAATGCTTACTCTGATGAAGTGTTCGGGAGCAACCACTGGTTTAAAGGAAAA 321
Qy 61 AGTAACCAAGAACTGTTCTCATTTGTTTAAAAAATGAAATGTTCAACTGAATGGAATA 120
Db 322 AGTAACCAAGAACTGTTCTCATTTGTTTAAAAAATGAGGATGTTCAACTGAATGGAATA 381
Qy 121 GATATCGGATGGAATAGTTTACAAAAAAGAGCTCGAGGAGCAGCTGAATCTTTACAA 180
Db 382 GATATCGGATGGAATAGTTTACAAAAAAGAGCTCGAGGAGCAGCTGAATCTTTACAA 441
Qy 181 CGAGGAGCGAAACTACTTGGGACCAAGAGGACATGGAATGGGAAACACACAGTGGAT 240
Db 442 CGAGGAGCGGAAACTACTTGGGACCAAGAGGACATGGAATGGGAAATCTACAGTGGAT 501
Qy 241 GAAATGACCAAAAAAGCAAGTATTTCATTTTGTGTTTAAAAAATGTTTATTTGAA 300
Db 502 GAAATGACCAAAAAAGCAAGTATTTCATTTTGTGTTTAAAAAATGTTTATTTGAA 561
Qy 301 GTGCTTAAACAAAGAATATATTTCTCTGGTATGTTAATTTGGTTGTGCAACATGAATGG 360
Db 562 GTGCTTAAACAAAGAATATATTTCTCTGGTATGTTAATTTGGTTGTGCAACATGAATGG 621
Qy 361 GGAAGAGACCAAGGCTGGCACTGCCATTAATTTGGAGAAAGAGCTTTTAGTCAAGCT 420
Db 622 GGAAGAGACCAAGGCTGGCACTGCCATTAATTTGGAGAAAGAGCTTTTAGTCAAGCT 681
Qy 421 CAAGGAAATGTTGGAGAAAGCACTAAATTTTACTTGGAGCAGATGTTTGGTGAACGCC 480
Db 682 CAAGGAAATGTTGGAGAAAGCACTAAATTTTACTTGGAGCAGATGTTTGGTGAACGCC 741
Qy 481 TGTAATGTGCAACTAACACACGCTGAAAGAAATTAAGAGAAATAGCAGAAAGCAAT 540

OTKKEYSIKITLKLVELXKVTSPEDMMWOPDSYIEMMAOPGGENIKNTLEICTLTL
ARTKYADLILEKAEKSLNFSLPDTRCKIFAFHGWNYIKVCHALICVLNRQGGKR
NTVLFHPASTGKSIIAQAIAQAVGVNVCYNAANVFPNDCTNKNKLIWVEEAGNFGQ
OWNQFAISGOTIRIDORKGSKQIEPTVIMTNNENITVVKIGCEERPEHTPIRD
RMLNHLTHLPDGLVDKNQWPMICAWLVKNQYQSTWASVCAKKGKVPDWTENWAE
KPVPTINSLSARSPTTPKSPPLSNVALTPLASDLBDLDALEPWSWTENTPVAGTAE
TONTGSAKSCADGOLSPWBSIEDLRACFGAEPLKXDFSEPLND"
join(2290..2320,2403..4561)
/gene="VPI"
join(2290..2320,2403..4561)
/gene="VPI"
/codon_start=1
/product="capsid protein"
/protein_id="AA61406.1"
/db_xref="GI:525327"
/translation="MAPPAKRAKRWVPGYKYLPGNSLDQSETPNPSDAAKEHDE
AYDKYIKSGNVPYLYFSADORDIDOKADGWGVHVFYFTRKFAFAPRLASSSEP
GTSVSIAGRTKPPPAHIFINOAARKKRLAAQOORTOTMSDGAOSPPDSGVASAA
RVERAADPGSGSGSGGGVGVSTGSVDNQHRYFLSDGWVELTAYSTRVHLYNP
KSENYCRVRVHNTNRTAGNMAKDAHEQIWTPLSLIDSNWQFSDWOFICNN
MSHVNLSLQDELNVVIKTVTEONTGAERAKIYNDLTASMMVALDSNNILPYTPAT
DNQETLGFYFWKFTSPYRYFENCDSISVTYDQTSIVDTMANASGLSOFYIE
NTQRIOLLRTGDFEFTCTYFETPEIKLSHTWOSNRQLGPOPOITDLPTADNENATLV
TRGDRSGIQTQISGNDVTEATRVPAQVGFCDPHDNFETSRAGFPVVPVNAVTOGN
EHDANGSLRTYDKQGDWGSNNKERTFWDALSYDSGRWADRCFINATPFTSPPAL
NLTITNSDPIGNKTAIHYQNVFNSYGLTAFPPHAPIYPOGQIWDKELDLDEHPRUHA
QAPFVCKNNAPOGLLRLAPNLTDQDPSNLSLRVYTGTFPFWKGLTKAKLRPNA
TWNPVTVQVSAQYQNEVEYMSIHKWLPATATGNMQSIPLLSRPVARNTY"

ORIGIN

Query Match 94.1%; Score 1899; DB 14; Length 5144;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1944; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
QY 1 ATGCTGGAATGCTTACTCTGATGAAGTTTGGGACCAACCTGGTTTAAAGGAAAA 60
DB 265 ATGCTGGAATGCTTACTCTGATGAAGTTTGGGACCAACCTGGTTTAAAGGAAAA 324
QY 61 AGTAACGAGGAAGTGTCTCATTTGTTTTTAAAAATGTTCAACTGAAATGGAATA 120
DB 325 AGTAACGAGGAAGTGTCTCATTTGTTTTTAAAACTGAGGATGTTCAACTAAATGGAATA 384
QY 121 GATATCGGATGGAATAGTTTCAAAAAAGAGCTGCAGAGAGCAGCTGAAATCTTTACAA 180
DB 385 GATATCGGATGGAATATTAACAAAAAGAGCTGCAGAGAGCAGCTGAAATCTTTACAA 444
QY 181 CGAGGAGCGAACTACTTGGGACCAAGCAGACATGGAATGGGAAACACACAGTGGAT 240
DB 445 CGAGGAGCGAACTACTTGGGACCAAGCAGACATGGAATGGGAAATCTACAGTGGAT 504
QY 241 GAAATGACCAAAAAAGCAAGTATTCATTTTGTGTTTAAAAAATGTTTATTTGAA 300
DB 505 GAAATGACCAAAAAAGCAAGTATTCATTTTGTGTTTAAAAAATGTTTATTTGAA 564
QY 301 GTGCTTAAACAAAGAATAATTTCTGTTGATGTTTAAATTTGTTTGGCAACATGATGG 360
DB 565 GTGCTTAAACAAAGAATAATGCTGCTGATGTTTAAATTTGTTTGGCAGCATGATGG 624
QY 361 GGAAGAAGCAAGCTGGCTGCTGATGTTTAAATTTGTTTGGGAAAGGATTTAGTCAAGCT 420
DB 625 GGAAGAAGCAAGCTGGCTGCTGATGTTTAAATTTGTTTGGGAAAGGATTTAGTCAAGCT 684
QY 421 CAAGGGAATTTGTTGGGAAGGAGCAGCTAAATTTTGTGTTGAGCAGATGGTTGTTAAAGGCC 480
DB 685 CAAGGGAATTTGTTGGGAAGGAGCAGCTAAATTTTGTGTTGAGCAGATGGTTGTTAAAGGCC 744
QY 481 TGTAAATGTCAACTAACACAGCTGAAAGAATAATTAACCTAAGAGAAATAGCAGAGACAAT 540
DB 745 TGTAAATGTCAACTAACACAGCTGAAAGAATAATTAACCTAAGAGAAATAGCAGAGACAAT 804
QY 541 GAGTGGGTACTCTTACTTATTAAGCATTAAGCAACCAACCAAGCACTATACCAAGTGT 600
DB 805 GAGTGGGTACTCTTACTTATTAAGCATTAAGCAACCAACCAAGCACTATACCAAGTGT 864


```
QY 1681 TCAGAAACTGGCGGAGCCAAAGTCCCACTCTTAAATTTACTAGTTCGGCAGC 1740
Db 1945 ACAGAAACTGGCGGAGCCGAGTCCCACTCTTAAATTTACTAGTTCGGCAGC 2004
QY 1741 TCACCAATTCAGACACCGAAAGTACGCTCTCAGCCAGCACTATGCACTAACTCCACTT 1800
Db 2005 TCACCAATTCAGACACCGAAAGTACGCTCTCAGCCAGCACTATGCACTAACTCCACTT 2064
QY 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGGAGCACACCAATATCTCTGTTCG 1860
Db 2065 GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGGAGCACACCAATATCTCTGTTCG 2124
QY 1861 GGCACTCGAGAAACCCAGAACACTGGGAAAGCTGGTTCCAAAGCCTGCCAAGATGTCAA 1920
Db 2125 GGCACTCGAGAAACCCAGAACACTGGGAAAGCTGGTTCCAAAGCCTGCCAAGATGTCAA 2184
QY 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGATTTGAGACGCTGCTCGTGGGAAACCG 1980
Db 2185 CTGAGCCCAACTTGGTCAGAGATCGAGGAGATTTGAGACGCTGCTCGTGGGAAACCG 2244
QY 1981 TTGAAGAAAGACTTCAGCGAGCGCTGAACCTTGGACTAA 2019
Db 2245 TTGAAGAAAGACTTCAGCGAGCGCTGAACCTTGGACTAA 2283

RESULT 12
MOU34254
LOCUS MOU34254 4764 bp DNA linear VRL 21-AUG-1996
DEFINITION Mouse parvovirus 1c DNA.
ACCESSION U34254
VERSION U34254.1 GI:1464794
KEYWORDS
SOURCE Mouse parvovirus 1c
ORGANISM Mouse parvovirus 1c
REFERENCE 1 (bases 1 to 4764)
AUTHORS Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L., Franklin,C.L., Hook,R.R. Jr. and Riley,L.K.
TITLE Molecular characterization of newly recognized rodent parvoviruses
JOURNAL J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
MEDLINE 9620434
PUBMED 8609486
REFERENCE 2 (bases 1 to 4764)
AUTHORS Besselsen,D.G.
TITLE Direct Submission
JOURNAL Submitted (17-AUG-1995) David G. Besselsen, Department of Veterinary Pathology, University of Missouri-Columbia, W213 Veterinary Medicine Building, Columbia, MO 65211, USA
FEATURES
source
1..4764
/organism="Mouse parvovirus 1c"
/mol_type="genomic DNA"
/specific_host="Mus musculus"
/db_xref="taxon:42842"

ORIGIN
Query Match 92.8%; Score 1873; DB 14; Length 4764;
Best Local Similarity 95.4%; Pred. No. 0;
Matches 1927; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 1 ATGCTGGAATGCTTACTCTGATGAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA 60
Db 122 ATGCTGGAATGCTTACTCTGATGAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA 181
QY 61 AGTAACAGGAGTGTCTCATTTGTTTTTAAATAAATGTTCAACTGAATGGAATA 120
Db 182 AGTAACAGGAGTGTCTCATTTGTTTTTAAATAAATGTTCAACTGAATGGAATA 241
QY 121 GATATCGGATGGAATAGTTTCAAAAAGAGCTCGAGGAGCGAGCTGGAATCTTTACAA 180
Db 242 GATATCGGATGGAATAGTTTCAAAAAGAGCTCGAGGAGCGAGCTGGAATCTTTACAA 301
QY 181 CGAGGAGCGGAAACTACTTGGGACCAAGCGAGGACATGGAATGGGAAACCAACAGTGGAT 240
```

```
Db 302 CGAGGAGCGGAAACTACTTGGGACCAAGCGAGGACATGGAATTCACAGTGGAT 361
QY 241 GAAATCACCACCAAAAGCAAGTATTCTTTTGTGTTTAAATAATGTTTATTGAA 300
Db 362 GAAGTGCACCAAAAGCAAGTATTCTTTTGTGTTTAAATAATGTTTATTGAA 421
QY 301 GTGCTTAAACACAAAGAAATATATTTCTGTGTGATTTAAATTTGTTTGTCAACATGAATGG 360
Db 422 GTGCTTAAACACAAAGAAATATATCTCTGTGTGATTTAAATTTGTTTGTGAGCATGAATGG 481
QY 361 GGAAGAAGACAGGCTTGGCACTGCCATGTACTTAATTTGGAGGAAGACATTTAGTCAAGCT 420
Db 482 GGAAGAAGACAGGCTTGGCACTGCCATGTACTTAATTTGGAGGAAGACATTTAGTCAAGCT 541
QY 421 CAAGGGAATTTGGTGGAGAAAGGCAACTAAATTTTCTGGAGCAGATGTTTGGTAAACAGCC 480
Db 542 CAAGGGAATTTGGTGGAGAAAGGCAACTAAATTTTCTGGAGCAGATGTTTGGTAAACAGCC 601
QY 481 TGTAAATGTGCAACTAACACACAGCTGAAAGAAATTTAAACTAAGAGAAATAGCAGAGAACAT 540
Db 602 TGTAAATGTACAGCTAACACACAGCTGAAAGAAATTTAAACTAAGAGAAATAGCAGAGACAGT 661
QY 541 GAGTGGTTTACTCTTACTTATTAAGCATAGCAAAACCAAAAGACATATACCAAGTGT 600
Db 662 GAATGGTTTACCTTGTCTCTTATTAAGCATAGCAAAACCAAAAGACATATACCAAGTGT 721
QY 601 GTTCTTTTGGAAACATGATTCCTTACTATTTTAACTAAAGAAATAAAGCACTAGT 660
Db 722 GTTCTTTTGGGCAACATATTTGCTTACTTTTTTAAACCAAGAAATAAAGCACTAGT 781
QY 661 CCACCAAGAGAGCGAGGCTATTTCTTAGCAGTGAATCTGGCTGGAAAACTTAACCTTTTAA 720
Db 782 CCGCCAAGGAGCGAGGCTATTTCTTAGTGAATCTGGCTGGAAAACTTAACCTTTTAA 841
QY 721 AAGAAGGCGAGCGCATCTAGTGAGCAAACTATATACATCTGATGACATGCGGCCAGAAACG 780
Db 842 AAGAAGGCGAGCGCATCTAGTGAGCAAACTATATACATCTGATGACATGCGGCCAGAAACG 901
QY 781 GTTGAACACACAGTAAACCACTGCGAGGAACTTAAGCGCGGAGAAATTCAAACTTAAATAA 840
Db 902 GTTGAACACACAGTAAACCACTGCGAGGAACTTAAGCGCGGAGAAATTCAAACTTAAATAA 961
QY 841 GAAGTTTCTTATTAATAACTACACTTAAAGAGCTGGTGCATATAAAGAGTAACCTCACCAGAG 900
Db 962 GAGTTTCTTATTAATAACCACTTAAAGAGCTAGTGCATATAAGAGAGTAACTCACCAGAA 1021
QY 901 GACTGATGATGATGAGCGCAGACATTTACATTTGAAATGATGCTCAACCCAGTGGAGAA 960
Db 1022 GACTGATGATGATGAGCGCAGACATTTATTTGAAATGATGCTCAACCCAGTGGAGAA 1081
QY 961 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAACCAAAACAGCA 1020
Db 1082 AACCTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCTAGAACCAAAACAGCA 1141
QY 1021 TTTGACTTAATTTTAAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCTGAC 1080
Db 1142 TTTGACTTGAATTTTAAAAAAGCTGAAACCCAGCAAACTAACCAACTTTTCACTGCCTGAC 1201
QY 1081 ACAAGAACCTGCAAGATTTTGTCTTTTATGCTTGAATCATGTTTAAAGTTTGCCTAGCT 1140
Db 1202 ACAAGAACCTGCAAGATTTTGTCTTTTATGCTTGAATCATGTTTAAAGTTTGCCTAGCT 1261
QY 1141 ATTTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATACTGTTTTTATTTTATGACCA 1200
Db 1262 ATTTGCTGTGTTTAAACAGACAGGAGGCAAAAGAAATACTGTTTTTATTTATGACCA 1321
QY 1201 GCCAGACAGGCAAAATCTATTATTGCAAGCAAGCAATAGCAGACAGTGTGGCAATTTGGT 1260
Db 1322 GCCAGTACAGGCAAAATCTATTATTGCAAGCAATAGCAGACAGTGTGGTAAATTTGGT 1381
QY 1261 TGTATAATGACGCAATGTAAACTTTTCATTTTATGATGTATCAACCAAGAACTTGATT 1320
```


Db 1382 TGTATTAATGACGAAATGTGAACCTTTCCATTCAATGACTGTACCAACAAGAACTTAATT 1441
Qy 1321 TGGGTAGAAGAGCTGTAATCTTTGGACAGCAAGTAACCAAGTTTAAAGCCATTGCTCT 1380
Db 1442 TGGGTGAAGAGCTGTAATCTTTGGACAAACAGTAACCAAGTTTAAAGCCATTGCTCT 1501
Qy 1381 GGTCAAGCTATTGCAATGATCAAAAGGAAAGGAGGAGCAACAGATTGAACCAACCA 1440
Db 1502 GGTCAAAACAATTGCAATGATCAAAAGGAAAGGAGGAGCAACAGATTGAACCAACCA 1561
Qy 1441 GTATCATGACCAACCAATGAGAACATTACAGTGTGTCAGAAATAGGCTCGGAAGAAAGACCA 1500
Db 1562 GTATCATGACCAACCAATGAGAACATTACAGTGTGTCAGAAATAGGCTCGGAAGAAAGACCA 1621
Qy 1501 GAACACACTCAACCAATCAGAGACAGAAATGCTTAAACATTATCATCTAACACATACATTG 1560
Db 1622 GAACACACTCAACCAATCAGAGACAGAAATGCTTAAACATTATCATCTAACACATACATTG 1681
Qy 1561 GGTGACTTTGGTTGGTTGCAAAATGAATGGCCCATGATTGCTGCTTGGTTGGTAAAG 1620
Db 1682 GGTGACTTTGGTTGGTTGCAAAATGAATGGCCCATGATTGCTGCTTGGTTGGTAAAG 1741
Qy 1621 AATGGTTACCAATCTACATGGCAAGCTACTGTGCTAAATGGGCAAAAGTTCTGTATTGG 1680
Db 1742 AATGGTTACCAATCTACATGGCAAGCTACTGTGCTAAATGGGCAAAAGTTCTGTATTGG 1801
Qy 1681 TCAGAAACTGGGCGAGGCAAAAGGTGCCAACTCTATAAATTTACTAGTTTGGCAGCG 1740
Db 1802 ACGGAAACTGGGCGAGGCGAAGGTGCCGACTCTCTATAAATTCACCTAGTTTGGCAGCG 1861
Qy 1741 TCACCAATTCAGACACCGGAAAGTACGCTCTGACGACAGAACTATGCACTAATCTCACTT 1800
Db 1862 TCACCAATTCAGACACCGGAAAGTACGCTCTGACGACAGAACTATGCACTAATCTCACTT 1921
Qy 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGAGGACACCAATATCTCTGTTGG 1860
Db 1922 GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGAGGACACCAATATCTCTGTTGG 1981
Qy 1861 GGCATCTCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCGCAAGATGGTCAA 1920
Db 1982 GGCATCTCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTGCGCAAGATGGTCAA 2041
Qy 1921 CTGAGCCCACTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTTCGGTGGGAACCG 1980
Db 2042 CTGAGCCCACTGGTCAGAGATCGAGGAGGATTTGAAAGCGTCTTCGGTGGGAACCG 2101
Qy 1981 TTGAAGAAAGACTTTCAGGAGCGCTGAACTTGACTAA 2019
Db 2102 TTGAAGAAAGACTTTCAGGAGCGCTGAACTTGACTAA 2140

RESULT 13
HOU34255
LOCUS
DEFINITION Hamster parvovirus DNA.
ACCESSION U34255
VERSION U34255.1 GI:1464792
KEYWORDS
SOURCE Hamster parvovirus
ORGANISM Hamster parvovirus
REFERENCE 1 (bases 1 to 4773)
Besselsen, D.G., Pintel, D.J., Purdy, G.A., Besch-Williford, C.L., Franklin, C.L., Hook, R.R. Jr. and Riley, L.K.
Molecular characterization of newly recognized rodent parvoviruses
J. Gen. Virol. 77 (Pt 5), 899-911 (1996)
96201434
MEDLINE
PUBMED 8609486
REFERENCE 2 (bases 1 to 4773)
Besselsen, D.G.
Direct Submission
Submitted (17-AUG-1995) David G. Besselsen, Department of
Veterinary Pathology, University of Missouri - Columbia, W213

FEATURES
source
Location/Qualifiers
1. 4773
/organism="Hamster parvovirus"
/mol_type="genomic DNA"
/specific_host="Mesocricetus auratus"
/db_xref="taxon:42843"

Veterinary Medicine Building, Columbia, MO 65211, USA

Query Match 92.2%; Score 1862.2; DB 14; Length 4773;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 1921; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

Qy 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAAA 60
Db 122 ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA 181
Qy 61 AGTAAACGAGGAAGTGTCTCATTTGTTTTAAAAATGAAATGTTCAACTGAATGGAAAA 120
Db 182 AGTAAACGAGGAAGTGTCTCATTTGTTTTAAAAATGAAATGTTCAAGTCAATGGAAAA 241
Qy 121 GATATCGGATGGAAATAGTTTACAAAAAAGAGCTGCAGGAGCAGAGCTGAAATCTTTACAA 180
Db 242 GATATCGGATGGAAATAGTTTACAAAAAAGAGCTGCAGGAGGAGAGCTGAAATCTTTACAA 301
Qy 181 CGAGGAGCGGAAACTACTTTGGGACCAAAAGCGAGGACATGGAATGGGAAACCCACAGTGGAT 240
Db 302 CGAGGAGCGGAAACTACTTTGGGACCGAGGAGGAGCATGGAATGGGAAATCTTCAGTGGAT 361
Qy 241 GAAATGACCAAAAAAGCAAGTATTCATTTTGAATCTTCTTGGTTAAAAAATGTTTATTTGAA 300
Db 362 GAAATGACCAAAAAAGCAAGTATTCATTTTGAATCTTCTTGGTTAAAAAATGTTTATTTGAA 421
Qy 301 GTCTTTAAACAACAAAGAAATATTTCTCTGGTGATGTTTAAATGGTTTGTGCAACATGAATGG 360
Db 422 GTCTGTGATCAAAAGAAACATAGCACCTAGTATGTTACTTGGTTTGTACAGCATGAATGG 481
Qy 361 GGAAGAGACCAAGCTGGCACTGCCATGTACTAATTTGGAGGAAAGGACTTTAGTCAAGCT 420
Db 482 GGAAGAGACCAAGCTGGCACTGTCTATGTAATTTGGAGGCAAGGACTTTAGCCAAGCT 541
Qy 421 CAGGGAATGTTGGAGAGGCACTAATATGTTTACTGGAGCAGATGTTGTTAAACAGCC 480
Db 542 CAGGGAATGTTGGAGAGGCACTAATATGTTTACTGGAGCAGATGTTGTTAAACAGCC 601
Qy 481 TGTAAATGTGCAACTTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGAACAT 540
Db 602 TGTAGTGTGAGCTATTACCAAGCTGAAAGAAATTAAGCTGAGAGAGATAGCGGAAGACAA 661
Qy 541 GAGTGGTTTACTCTTACTTACTTATAAGCAATAAGCAAAACCAAAAAAGACTATACCAAGTGT 600
Db 662 GAATGGTCACTTTGCTTACTTATAAGCAATAAGCAAAACCAAAAAAGACTATACCAAGTGT 721
Qy 601 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTTAAAGAAAAATTAAGCACTAGT 660
Db 722 GTTGTCTTGGAAATATAGTGTCTTACTTACTTTTATCCAAAGAGAAAAATATGACCAGT 781
Qy 661 CCACCAAGAGACGAGGCTATTTTCTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA 720
Db 782 CCACCAAGAGACGAGGCTATTTTCTAGCAGTACTCTGGCTGGAAAACTAACTTTTAA 841
Qy 721 AAAGAGGCGGAGCGCCATCTAGTGAGCAAACTATACATCTGATGACATCGCGCCAGAACG 780
Db 842 AAAGAGGCGGAGCGCCATCTAGTGAGCAAACTATACATCTGATGACATCGCGCCAGAACG 901
Qy 781 GTTGAACCAACAGTAACCACTGGCAGGAAACTAAGCGGCGAGAAATTCAACTTAAAAA 840
Db 902 GTTGAACCAACAGTAACCACTGGCAGGAAACTAAGCGGCGAGAAATTCAACTTAAAAA 961
Qy 841 GAAGTTTCTATTAAAACTACACTTAAAGAGCTGTCATAAAAAGAGTAACTCACCAGAG 900
Db 962 GAGGTCTTATTAAAAACCACTTAAAGAGCTGTCATTAAGAGAGTAACTCACCAGAA 1021

901 GACTGGATGATGTCAGCAGACAGTTCATTGAAATGATGCTCAACACAGGTGAGAA 960
1022 GACTGGATGATGTCAGCAGACAGTTCATTGAAATGATGCTCAACACAGGTGAGAA 1081
961 AACTGCTGAAAATACGCTAGAGATTGTACACTAACTCTAGCCAGAACCAACAGCA 1020
1082 AACCTGCTGAAAATACGCTAGAGATTGTACACTAACTCTAGCCAGAACCAACAGCA 1141
1021 TTTGACTTAATTTAGAAAAGCTGAAACCCAGCAAACTAACCACTTTTCACTGCTGAC 1080
1142 TTTGACTTAATTTAGAAAAGCTGAAACCCAGCAAACTAACCACTTTTCACTGCTGAC 1201
1081 ACAAGAACCTGCAGAAATTTTGGCTTTTCATGGCTGAACATATGTTAAAGTTGCCATGCT 1140
1202 ACAAGAACCTGCAGAAATTTTGGCTTTTCATGGCTGAACATATGTTAAAGTTGCCATGCT 1261
1141 ATTGCTGTGTTTAAACAGACAGAGGCAAGAAAAGAAATACCTGTTTATTTATGACCA 1200
1262 ATTGCTGTGTTTAAACAGACAGAGGCAAGAAAAGAAATACCTGTTTATTTATGACCA 1321
1201 GCCAGACAGGCAAACTATTATTGACAGCCATAGCACAAGCAGTTGGCAATGTTGGT 1260
1322 GCCAGTACAGCAATCCATATTGACAGCCATAGCACAAGCAGTTGGTATGTTGGT 1381
1261 TGCTATAATCCAGCCAAATGTAAACTTTCCATTTAATGACTGTACCAACAGAACTTGAT 1320
1382 TGCTATAATCCAGCAATGTAAACTTTCCATTTAATGACTGTACCAACAGAACTTGAT 1441
1321 TGGGTAGAAAGCTGGTAACTTTGGACAGCAAGTAAACAGTTTAAAGCAATTTGCTCT 1380
1442 TGGGTAGAAAGCTGGTAACTTTGGACAGCAAGTAAACAGTTTAAAGCAATTTGCTCT 1501
1381 GGTCAAGCTATTGCTGATTAATCAAAAGGAAAGGAGCAAGAAAGTGAACCAACCA 1440
1502 GGTCAAGCTATTGCTGATTAATCAAAAGGAAAGGAGCAAGAAAGTGAACCAACCA 1561
1441 GTCATCATGACCAAAATAGAACATTACAGTGCTGAGAAATAGGCTGCGAAGAAAGCA 1500
1562 GTCATCATGACCAAAATAGAACATTACAGTGCTGAGAAATAGGCTGCGAAGAAAGCA 1621
1501 GAACACACTCAACCAATCAGACAGAGATGCTTAACATTTCACTAACACATACCTGCT 1560
1622 GAACACACTCAACCAATCAGACAGAGATGCTTAACATTTCACTAACACATACCTGCT 1681
1561 GGTGACTTTGGTTGGTTGACAAAATGAAATGGCCCAATGATTTGCTGCTGGTTGGTAAAG 1620
1682 GGTGACTTTGGTTGGTTGACAAAATGAAATGGCCCAATGATTTGCTGCTGGTTGGTAAAG 1741
1621 AATGGTTACCAATCTACATGGCAAGCTACTGTGCTAAATGGGGCAAGTTCTGTATGG 1680
1742 AATGGTTACCAATCTACATGGCAAGCTACTGTGCTAAATGGGGCAAGTTCTGTATGG 1801
1681 TCAGAAACTGGGGAGCCAGAGTGGCAAGTGGCTTAAATTTACTAGTTGCGGACGC 1740
1802 ACGAAACTGGGGAGCCAGAGTGGCAAGTGGCTTAAATTTACTAGTTGCGGACGC 1861
1741 TCACCAATTCAGCAACCGAAAGTACGCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800
1862 TCACCAATTCAGCAACCGAAAGTACGCTCTCAGCCAGAACTATGCACTAACTCCACTT 1921
1801 GCATCGGATCTGAGGACCTGGCTTTAGAGCTTTGAGGACACACCAATATCTCTGTTGGC 1860
1922 GCATCGGATCTGAGGACCTGGCTTTAGAGCTTTGAGGACACACCAATATCTCTGTTGGC 1981
1861 GGCACTCAGAAACCCAGAACACTGGGAGAGCTGTTCCAAAGCTCGCAAGATGGTCAA 1920
1982 GGCACTCAGAAACCCAGAACACTGGGAGAGCTGTTCCAAAGCTCGCAAGATGGTCAA 2041
1921 CTGAGCCCAACTTGGTCAGAGATCGAGAGGATTTGAGAGCGTCTTCTGGTGGCAACCG 1980
2042 CTGAGCCCAACTTGGTCAGAGATCGAGAGGATTTGAGAGCGTCTTCTGGTGGCAACCG 2101
1981 TTGAAGAAAGACTTCAGCGAGCCGCTGAATTTGGACTAA 2019

2102 TTGAAGAGAGACTTCAGCGAGCCGCTGAATTTGGACTAA 2140
RESULT 14
MVU34256
LOCUS
DEFINITION
Mice minute virus DNA.
ACCESSION
U34256
VERSION
U34256.1
KEYWORDS
GI:1464795
SOURCE
Mice minute virus
ORGANISM
Mice minute virus
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
REFERENCE
1 (bases 1 to 4761)
Besselsen,D.G., Pintel,D.J., Purdy,G.A., Besch-Williford,C.L.,
Franklin,C.L., Hook,R.R. Jr. and Riley,L.K.
Molecular characterization of newly recognized rodent parvoviruses
J. gen. Virol. 77 (Pt 5), 899-911 (1996)
AUTHORS
J. gen. Virol. 77 (Pt 5), 899-911 (1996)
TITLE
Molecular characterization of newly recognized rodent parvoviruses
JOURNAL
J. gen. Virol. 77 (Pt 5), 899-911 (1996)
MEDLINE
96201434
PUBMED
8609486
REFERENCE
2 (bases 1 to 4761)
Besselsen,D.G.
AUTHORS
Besselsen,D.G.
TITLE
Direct Submission
JOURNAL
Submitted (17-AUG-1995) David G. Besselsen, Department of
Veterinary Pathology, University of Missouri-Columbia, W213
Veterinary Medicine Building, Columbia, MO 65211, USA
LOCATION/Qualifiers
FEATURES
source
1. .4761
/organism="Mice minute virus"
/mol_type="genomic DNA"
/strain="Cutter"
/specific_host="Mesocricetus auratus"
/db_xref="taxon:10794"
ORIGIN
Query Match 91.9%; Score 1855.8; DB 14; Length 4761;
Best Local Similarity 94.9%; Pred. No. 0;
Matches 1917; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
QY 1 ATGCTCGGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 60
DB 122 ATGCTCGGAAATGCTTACTCTGATGAGTTTGGGAGCAACCACTGGTTAAAGGAAAA 181
QY 61 AGTAAACAGAGAGTGTCTCATTTGTTTAAATAAGTTCATCACTGATGAGAA 120
DB 182 AGTAAACAGAGTGTCTCATTTGTTTAAATAAGTTCATCACTGATGAGAA 241
QY 121 GATATCGGATGGAATAGTTTACAAAAGAGCTCAGAGGAGCAGCTGAAATCTTTACAA 180
DB 242 GATATCGGATGGAATAGTTTACAAAAGAGCTCAGAGGAGCAGCTGAAATCTTTACAA 301
QY 181 CGAGGAGCGGAAACTACTTGGGACCAAGCAGGACATGGAATGGGAAACCACTGATGAT 240
DB 302 CGAGGAGCGGAAACTACTTGGGACCAAGCAGGACATGGAATGGGAAACCACTGATGAT 361
QY 241 GAAATGACCAAAAGCAAGTATTCATTTTGTATCTTTGTTTAAATAAGTTCATTTTAA 300
DB 362 GAACTAAACCAAAAGCAAGTATTCATTTTGTATCTTTGTTTAAATAAGTTCATTTTAA 421
QY 301 GTCCTTAAACCAAAAGCAAGTATTCATTTTGTATCTTTGTTTAAATAAGTTCATTTTAA 360
DB 422 GTGCTTAAGTACAAAGCAAGTATTCATTTTGTATCTTTGTTTAAATAAGTTCATTTTAA 481
QY 361 GGAAGAGCAAGCTGGCACTGCCATGTACTAATTTGGAGAAAGGACTTTAGTCAAGCT 420
DB 482 GGAAGAGCAAGCTGGCACTGCCATGTACTAATTTGGAGAAAGGACTTTAGTCAAGCT 541
QY 421 CAAGGAAATGCTGGAGAGGCAAGTAAATTTTACTTGGAGCAGATGTTGTTAAACAGCC 480
DB 542 CAAGGAAATGCTGGAGAGGCAAGTAAATTTTACTTGGAGCAGATGTTGTTAAACAGCC 601
QY 481 TGTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAAAATAGCAGAGCAAT 540

```
Db 602 TGAGTGTGCGCTATCACCAGCGGAAAGAAATTAAGCTGAGAGAAATAGCGGAGAACCA 661
Qy 541 GAGTGGGTACTCTACTTACTTATAAGCATAGCAAAACCAAAAGAGCTATACCAAGTGT 600
Db 662 GAATGGTCACTTGTCTTACTTATAAGCATAGCAAAACCAAAAGAGCTATACCAATGT 721
Qy 601 GTTCTTTTGGAAACATGATGCTTACTTATTTTAACTAAAGAAAATAAGCACTAGT 660
Db 722 GTTGTCTTGGAAATATGATGCTTACTTATTTTAACTAAAGAAAATAAGCACTAGT 781
Qy 661 CCACCAAGACAGCGAGCTATTTCTTAGCAGTGAAGTCTCTGCTGGAAATCTAACTTTT 720
Db 782 CCACCAAGGACGAGGCTATTTCTTAGCAGTGAAGTCTCTGCTGGAAATCTAACTTTT 841
Qy 721 AAAGAAAGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG 780
Db 842 AAAGAAAGCGAACGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCCAGAAACG 901
Qy 781 GTTGAACACACAGTAACCACTGGCGAGGAACTAAGCGCGGCGAGAACTTCAAACTAA 840
Db 902 GTTGAACACACAGTAACCACTGGCGAGGAACTAAGCGCGGCGAGAACTTCAAACTAA 961
Qy 841 GAAGTTTCTATTAAACCTACACTTAAAGAGCTGTGCATAAAGAGTAACCTCACAGAG 900
Db 962 GAGTTTCTATTAAACCACTTAAAGAGCTGTGCATAAAGAGTAACCTCACAGAA 1021
Qy 901 GACTGGATGATGACGAGCAGACAGTTACATTTGAAATGATGCTCAACAGAGTGGAGAA 960
Db 1022 GACTGGATGATGACGAGCAGACAGTTACATTTGAAATGATGCTCAACAGAGTGGAGAA 1081
Qy 961 AACTGCTGAAAAATACGCTAGAGATTTGTGACATACTAAGCGCGGCGAGAACTTCAAACTAA 1020
Db 1082 AACTGCTGAAAAATACGCTAGAGATCTGTACGCTAAGCTAGCTAGTAAAGTTTGGCATGCT 1141
Qy 1021 TTTGACTTAATTTAGAAAGCTGAAACAGCAAACTAACCACTTTTCACTGCCTGAC 1080
Db 1142 TTTGACTTGAATTTAGAAAGCTGAAACAGCAAACTAACCACTTTTCACTGCCTGAC 1201
Qy 1081 ACAAGAACTGCAAGATTTTGTCTTTTATGCTGGAATCTATGTTAAAGTTTGGCATGCT 1140
Db 1202 ACAAGAACTGTAAGATTTTGTCTTTTATGCTGGAATCTATGTTAAAGTTTGGCATGCT 1261
Qy 1141 ATTTGCTGTTTAAACAGAGGAGCAAAAGAAATCTGTTTTATTTATGAGACCA 1200
Db 1262 ATTTGCTGTTTAAACAGAGGAGCAAAAGAAATCTGTTTTATTTATGAGACCA 1321
Qy 1201 GCCAGACAGGCAATCTATTATGCAAGCCATAGCAAGCACTGCGCAATGTTGT 1260
Db 1322 GCCAGTACAGGCAATCCATCTATGCAAGCCATAGCAAGCACTGCGCAATGTTGT 1381
Qy 1261 TGCTATAATGAGCAATGTAACTTTTCCATTTAATGACTGTACCAAGCAAGCACTGATT 1320
Db 1382 TGCTATAATGAGCAATGTAACTTTTCCATTTAATGACTGTACCAAGCAAGCACTGATT 1441
Qy 1321 TGGGTAGAAGAGCTGTAATCTTTGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT 1380
Db 1442 TGGGTAGAAGAGCTGTAATCTTTGACAGCAAGTAAACAGTTTAAAGCCATTTGCTCT 1501
Qy 1381 GGTCAAGCTATTGCAATGATCAAAAGGAAAGGAGCAAAAGAGTGAACCAACCA 1440
Db 1502 GGTCAAGCTATTGCAATGATCAAAAGGAAAGGAGCAAAAGAGTGAACCAACCA 1561
Qy 1441 GTCATATGACCAAAATGAGCAATTAAGTGTGAGAAATGAGTGTGCGAAGAAAGACCA 1500
Db 1562 GTCATATGACCAAAATGAGCAATTAAGTGTGAGAAATGAGTGTGCGAAGAAAGACCA 1621
Qy 1501 GAAACACTCAACCAATCAGAGACAGAAATCTTAACATTTCACTAACAATCCTTGCCT 1560
Db 1622 GAAACACTCAACCAATCAGAGACAGAAATCTTAACATTTCACTAACAATCCTTGCCT 1681
Qy 1561 GGTGACTTTGTTGTTGAAAAATGAATGGCCCATGATTTGTGCTGTTGGTAAAG 1620
```

```
Db 1682 GGTGACTTTGTTGTTGACAAAAATGAATGCCCATGATTTGTGCTGTTGTTAAAG 1741
Qy 1621 AATGGTTACCAATCTACATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTGATTGG 1680
Db 1742 AATGGTTACCAATCTACATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTCTGATTGG 1801
Qy 1681 TCAGAAAACTGGCGGAGCCAAAGGTGCCAACTCTATAAATTTACTAGGTTCGGACGC 1740
Db 1802 TCAGAAAACTGGCGGAGCCAAAGGTGCCAACTCTATAAATTTACTAGGTTCGGACGC 1861
Qy 1741 TCACCAATTCACGACACCGAAAAAGTACGCCCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800
Db 1862 TCACCAATTCAGGACACCGAAAAAGTACGCCCTCTCAGCCAGAACTATGCACTAACTCCACTT 1921
Qy 1801 GCATCGATCTCAGGACCTGGCTTTAGAGCCTTTGAGAGCCTTGAGAGCACACAAATATCTCTGTTGCG 1860
Db 1922 GCATCGATCTCAGGACCTGGCTTTAGAGCCTTTGAGAGCCTTGAGAGCACACAAATATCTCTGTTGCG 1981
Qy 1861 GGCACCTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920
Db 1982 GGCACCTGCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 2041
Qy 1921 CTGAGCCCACTTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTCGGTGGGAAACCG 1980
Db 2042 CTGAGCCCACTTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTCGGTGGGAAACCG 2101
Qy 1981 TTGAAGAAAGACTTTCAGGAGCGCTGAACTTGGACTAA 2019
Db 2102 TTGAAGAAAGACTTTCAGGAGCGCTGAACTTGGACTAA 2140

RESULT 15
LOCUS CO786765
DEFINITION Sequence 1 from Patent WO2004018689.
ACCESSION CO786765
VERSION CO786765.1 GI:45721778
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Iggo,R. and Mallerba,M.
TITLE Anti-neoplastic viruses
JOURNAL Patent: WO 2004018689-A 1 04-MAR-2004;
BTG INTERNATIONAL LIMITED (GB)
FEATURES
source 1: 5121
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="Parvovirus H1\WM (pH1) - Parvovirus H1 with
promoter P4 and left hairpin from WM"

ORIGIN
Query Match 87.7%; Score 1771; DB 6; Length 5121;
Best Local Similarity 92.3%; Pred. No. 0;
Matches 1864; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

Qy 1 ATGCGCTGGAATCTTACTCTGATGAAGTTTTCGGAGCAACCACTGTTTAAAGAAAAA 60
Db 282 ATGCGCTGGAATCTTACTCTGATGAAGTTTTCGGAGCAACCACTGTTTAAAGAAAAA 341
Qy 61 AGTAAACAGGAAGTGTCTCATTTGTTTAAATAAGTGAATGTTCAACTGAATGAAAA 120
Db 342 AGTAAACAGGAAGTGTCTCATTTGTTTAAATAAGTGAATGTTCAACTGAATGAAAA 401
Qy 121 GATATCGGATGGAATAGTTACAAAAAGAGCTCAGAGGACGAGCTGAAATCTTTTCAA 180
Db 402 GATATCGGATGGAATAGTTACAAAAAGAGCTCAGAGGACGAGCTGAAATCTTTTCAA 461
Qy 181 CGAGGACGGAATCTACTTGGGACCAAGCAGGACATGGAATGGGAAACCACTGAT 240
```

Db 462 CGAGGAGCGGAAACTACTTTGGACCAACGAGGACATGGAATGGGAAACCAAGTGGAT 521
Qy 241 GAAATGACCAAAAGCAAGTATTCATTTTGGTTTAAATAATGTTTATTGAA 300
Db 522 GAAATGACCAAAAGCAAGTATTCATTTTGGTTTAAATAATGTTTATTGAA 581
Qy 301 GTGCTTAAACAACAAGAAATATTTCTGTTGATGTTAAATGTTTGTGCAACATGAATGG 360
Db 582 GTGCTTAAACAACAAGAAATATTTCTGTTGATGTTAAATGTTTGTGCAACATGAATGG 641
Qy 361 GGAAGACCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAAGACATTTAGTCAAGCT 420
Db 642 GGAAGACCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAAGACATTTAGTCAAGCT 701
Qy 421 CAAGGGAATTTGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTGGTAAACAGCC 480
Db 702 CAAGGGAATTTGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTGGTAAACAGCC 761
Qy 481 TGTAAATGTGCACTAACACAGCTGAAGAAATTTAACTAAGAGAAATAGCAGAGACAAT 540
Db 762 TGTAAATGTGCACTAACACAGCTGAAGAAATTTAACTAAGAGAAATAGCAGAGACAAT 821
Qy 541 GAGTGGTGTACTCTACTTACTTATAGCATAAGCAACCAAAAAGACATATACCAAGTGT 600
Db 822 GAGTGGTGTACTCTACTTACTTATAGCATAAGCAACCAAAAAGACATATACCAAGTGT 881
Qy 601 GTTCTTTTGGAAACATGATGTTCTTATTTTAACTAATAAAGAAATAGCACTAGT 660
Db 882 GTTCTTTTGGAAACATGATGTTCTTATTTTAACTAATAAAGAAATAGCACTAGT 941
Qy 661 CCACCAAGACGAGGCTATTTCTTAGCAGTACTCTGGCTGGAAACTTAACTTTTAA 720
Db 942 CCACCAAGACGAGGCTATTTCTTAGCAGTACTCTGGCTGGAAACTTAACTTTTAA 1001
Qy 721 AAAGAAGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATGCGGCGAGAAACG 780
Db 1002 AAAGAAGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGATGAAACCAAGAACG 1061
Qy 781 GTTGAACCAAGTAAACCACTGCCAGGAACTAAGCGCGGAGAAATTCAACTTAAATA 840
Db 1062 GTCGAGACCAAGTGAACCACTGCCAGGAACTAAGCGCGGAGAAATTCAACTTAAATA 1121
Qy 841 GAACTTTCTATTTAAACTTACACTTAAAGAGCTGGTGCATATAAAGAGTAACTCACAGAG 900
Db 1122 GAGGTCTCGATTAAACCACTCAAGAGTGGTGATATAAAGAGTAACTCACAGAG 1181
Qy 901 GACTGGATGATGACGACGACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Db 1182 GACTGGATGATGACGACGACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1241
Qy 961 AACCTGCTGAAATAACGCTAGAGATTTGTACACTTAAGCTAGCCAGAACCAAAACAGCA 1020
Db 1242 AACCTGCTGAAATAACGCTAGAGATTTGTACACTTAAGCTAGCCAGAACCAAAACAGCC 1301
Qy 1021 TTTGACTTTAATTTAGAAAAAGCTGAAACCAAGCAAACTAAACAACTTTTCTGCTGAC 1080
Db 1302 TTTGACTTTAATTTAGAAAAAGCTGAAACCAAGCAAACTAGCCAACTTTTCTGCTGAC 1361
Qy 1081 ACAAGAACTGACAGAAATTTTGTCTTTTCTGCTGGAATCTGTTTAAAGTTTGGCATGCT 1140
Db 1362 ACCAGAACCTGTAGAAATTTTGTCTGAGCATGCTGGAATCTGTTTAAAGTTTGGCATGCT 1421
Qy 1141 ATTTGCTGTGTTTAAACAGACAGAGGCGAAAGAAATCTGTTTAAATTTCTGAGGACCA 1200
Db 1422 ATCTGTGTGTCTGAAATAGACAGAGGCGAAAGAAACACTGTGCTCTTTCAAGGACCA 1481
Qy 1201 GCCAGACAGGCAATCTATTATTGCAAGCCATAGCAGCAGTGGCAATGTTGCT 1260
Db 1482 GCCAGACAGGCAATCTATTATTGCAAGCCATAGCAGCAGTGGCAATGTTGCT 1541
Qy 1261 TGCTAATGCGCCCAATCTAACTTTTCCATTTTAAATGACTGTACCAACAGAACTTGATT 1320
Db 1542 TGTTAATGCTGCCAATGTAATGAACTTTTCCATTTTAAATGACTGTACCAACAGAACTTGATT 1601

Qy 1321 TGGCTAGAGAAGCTGGTAACTTTGGACAGCAAGTAAACCAAGTTTAAAGCAATTTGCTCT 1380
Db 1602 TGGCTAGAGAAGCTGGTAACTTTGGCCAGCAAGTAAACCAATTTCAAAGCTATTTGTTCT 1661
Qy 1381 GGTCAAGCTATTTCCATTTGATCAAAAAAGGAGGCAAAACAGATTGAACCAACACCA 1440
Db 1662 GGCCAAACCATACGATTTGATCAAAAAAGGAGGCAAAACAGATTGAACCAACACCA 1721
Qy 1441 GTCATCATGACCACAAAATGAGAACATTTACAGTGTGTCAGATAGGCTGCGAAGAAAGACCA 1500
Db 1722 GTTATTATGACCAACCAAGCAATTTACCGTGTGTAGATAGGCTGTGAGGAAAGACCA 1781
Qy 1501 GAAACACTCAACCAATCAGAGACAGAAATGCTTAACTTCAATCTAAACATGACGACGATACCT 1560
Db 1782 GAAACACTCAACCAATCAGAGACAGAAATGCTTAACTTCAATCTAAACATGACGACGATACCT 1841
Qy 1561 GGTGACTTTGGTTGGTTGACAAAAATGAATGGCCCATGATTTGCTTGGTTGGTAAAG 1620
Db 1842 GGTGACTTTGGTTGGTTGATGAACCAAGATGCTTGAATGGCTCTGATCTGTGCTTGGTTGGTGAAG 1901
Qy 1621 AATGGTTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGCAAAAGTTTCTGATTTGG 1680
Db 1902 AATGGTTTACCAATCTACCATGGCTTGTACTGTGCTAAATGGGCAAAAGTTTCTGATTTGG 1961
Qy 1681 TCAGAAAACTGGGCGAGCCAAAAGGTGCAACTCTCTATAAATTTACTAGTTTCGGCAGCC 1740
Db 1962 TCAGAGGACTGGGCGAGCCGAAAGCTAGACACTCTCTATAAATTTGCTAGGTTCAATGGC 2021
Qy 1741 TCACCAATTCAGACACCGAAGTACGCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800
Db 2022 TCACCAATCTGACTCCGAGAAATGAGCTCTCAGCCAAACTAGCTCTTACTCCACTT 2081
Qy 1801 GCATCGGATCTCGAGGACCTGGCTTTTAGAGCTTTGGAGCCACCAAAATCTCCTGTTGG 1860
Db 2082 GCATCGGATCTCGAGGACCTGGCTTTTAGAGCTTTGGAGCCACCAAAATCTCCTGTTGG 2141
Qy 1861 GGCACTGAGAAAAACCAAGAACCACTGGGGAAGCTGTTCCAAAGCCTGCCAAGATGTTCAA 1920
Db 2142 GGCACTGAGAAAAACCAAGAACCACTGGGGAAGCTGTTCCAAAGCCTGCCAAGATGTTCAA 2201
Qy 1921 CTGAGCCCAACTTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTGGTGGGAACG 1980
Db 2202 CGGAGCCCAACTTTGGTTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTGGTGGGAACG 2041
Qy 1981 TTGAAGAAAGACTTCAGCGAGCCGCTGAACCTTGGACTAA 2019
Db 2262 TTGGAGGCGACTTTCACGAGGAGCTGACCTTGGACTAA 2300

Search completed: January 22, 2005, 21:19:34
Job time : 8548.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2005, 23:46:57 ; Search time 950.5 Seconds
(without alignments)
11150.532 Million cell updates/sec

Title: US-10-069-056-16
Perfect score: 2019
Sequence: 1 atggcgggaatgcttactc.....agccgtgaacttggaactaa 2019

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_23Sep04:.*
1: Geneseqn1980s:.*
2: Geneseqn1990s:.*
3: Geneseqn2000s:.*
4: Geneseqn2001as:.*
5: Geneseqn2001bs:.*
6: Geneseqn2002as:.*
7: Geneseqn2002bs:.*
8: Geneseqn2003as:.*
9: Geneseqn2003bs:.*
10: Geneseqn2003cs:.*
11: Geneseqn2003ds:.*
12: Geneseqn2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2019	100.0	2019	5	AAD02805 Parvoviru
2	2017.4	99.9	2019	5	AAD02797 Parvoviru
3	2017.4	99.9	5149	12	Adg39767 Minute vi.
4	2015.8	99.8	2019	5	AAD02801 Parvoviru
5	2015.8	99.8	2019	5	AAD02803 Parvoviru
6	2015.8	99.8	2019	5	AAD02799 Parvoviru
7	1040.4	51.5	5049	2	Aat15311 Non-atten
8	1040.4	51.5	5049	2	Aat15312 Attenuate
9	1040.4	51.5	5049	2	Aat88324 Attenuate
10	1038.8	51.5	5049	2	Aat88321 Attenuate
11	1038.8	51.5	5049	2	Aat88320 Canine pa
12	901	44.6	3524	1	Aan40252 Sequence
13	307.4	15.2	374	6	Abq95626 Tumour su
14	285.8	14.2	421	6	Abq94779 Tumour su
15	284.8	14.1	423	6	Abq94724 Tumour su
16	284.8	14.1	464	6	Abq94719 Tumour su
17	284.8	14.1	473	6	Abq94716 Tumour su
18	284.8	14.1	486	6	Abq94707 Tumour su
19	284.8	14.1	491	6	Abq94777 Tumour su
20	284.2	14.1	420	6	Abq94787 Tumour su
21	283.8	14.1	343	6	Abq94737 Tumour su

22	283.8	14.1	343	6	ABQ94739	Abq94739 Tumour su
23	283.8	14.1	403	6	ABQ94782	Abq94782 Tumour su
24	283.2	14.0	497	6	ABQ94718	Abq94718 Tumour su
25	280.4	13.9	340	6	ABQ94732	Abq94732 Tumour su
26	278.2	13.8	342	6	ABQ94765	Abq94765 Tumour su
27	273.8	13.6	420	6	ABQ94781	Abq94781 Tumour su
28	269.6	13.4	324	6	ABQ94740	Abq94740 Tumour su
29	263.4	13.0	339	6	ABQ95625	Abq95625 Tumour su
30	261.6	13.0	424	6	ABQ94778	Abq94778 Tumour su
31	163.8	8.1	468	6	ABQ94873	Abq94873 Tumour su
32	162.2	8.0	451	6	ABQ94868	Abq94868 Tumour su
33	162.2	8.0	457	6	ABQ94867	Abq94867 Tumour su
34	162.2	8.0	465	6	ABQ94874	Abq94874 Tumour su
35	162.2	8.0	469	6	ABQ94869	Abq94869 Tumour su
36	162.2	8.0	472	6	ABQ94880	Abq94880 Tumour su
37	162.2	8.0	515	6	ABQ94793	Abq94793 Tumour su
38	162.2	8.0	516	6	ABQ94807	Abq94807 Tumour su
39	162.2	8.0	516	6	ABQ94794	Abq94794 Tumour su
40	162.2	8.0	516	6	ABQ94800	Abq94800 Tumour su
41	162.2	8.0	517	6	ABQ94811	Abq94811 Tumour su
42	162.2	8.0	530	6	ABQ94791	Abq94791 Tumour su
43	162.2	8.0	530	6	ABQ94799	Abq94799 Tumour su
44	161	8.0	516	6	ABQ94796	Abq94796 Tumour su
45	160.6	8.0	530	6	ABQ94805	Abq94805 Tumour su

ALIGNMENTS

RESULT 1
AAD02805
ID AAD02805 standard; DNA; 2019 BP.
XX
AC AAD02805;
XX

DT 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX

DE Parvovirus non-structure protein 1 (NS1) variant (T463A) DNA.

XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoral disease; gene therapy; mutant; mutein; variant; ds.

XX Parvovirus.
OS Synthetic.

FH Key Location/Qualifiers
FT CDS 1..2019
FT /*tag= a
FT /product= "NS1 variant (T463A) protein"
FT mutation replace(1387, A)
FT /*tag= b

XX EPI077260-A1.

XX 21-FEB-2001.

XX 13-AUG-1999; 99EP-00115161.

XX 13-AUG-1999; 99EP-00115161.

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX Nueesch J, Rommelaere J;

XX WPI; 2001-212717/22.

XX P-PSDB; AAY72710.

XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.

XX Claim 7; Page 27-30; 4lpp; English.

XX	The present sequence is a DNA encoding parvovirus non-structure protein 1
CC	(NS1) variant (T463A). The invention relates to the variants of the
CC	parvovirus non-structure protein (NS1) having a shifted equilibrium
CC	between the DNA replication and transcription activities, and the
CC	cytotoxicity activity. These variants are useful as toxins for treating
CC	tumoural diseases. The variant DNAs are useful as vectors for gene
CC	therapy. (Updated on 06-AUG-2003 to correct OS field.)
XX	
SQ	Sequence 2019 BP; 697 A; 389 C; 453 G; 480 T; 0 U; 0 Other;
	Query Match 100.0%; Score 2019; DB 5; Length 2019;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 2019; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ATGCCTGGAAATGCTTACTCTCATGAAGTTTTGGGAGCAACCAACTGGTTAAAGGAAAAA 60
DB	1 ATGCCTGGAAATGCTTACTCTCATGAAGTTTTGGGAGCAACCAACTGGTTAAAGGAAAAA 60
QY	61 AGTAACACAGGAAGTGTTTCATTGTGTTTTTAAAATAAGAAATGTTCAAATGGAAAAA 120
DB	61 AGTAACACAGGAAGTGTTCTATTGTGTTTTTAAAATAAGAAATGTTCAAATGGAAAAA 120
QY	121 GATATCGGATGGNATAGTTACAABAAAGAGCTGCGAGGAGCGAGCTGAATCTTTACNA 180
DB	121 GATATCGGATGGNATAGTTACAABAAAGAGCTGCGAGGAGCGAGCTGAAATCTTTACAA 180
QY	181 CGAGGAGCGMAACTACTTTGGGACCAAGCAGGACATGGAAATGGAAACCAACAGTGGAT 240
DB	181 CGAGGAGCGMAACTACTTTGGGACCAAGCAGGACATGGAAATGGGAAACCAACAGTGGAT 240
QY	241 GAAATGACCACAAAAGCAAGTATTCATTTTGTATCTTTGGTTAAAAAATGTTTATTTGAA 300
DB	241 GAAATGACCACAAAAGCAAGTATTCATTTTGTATCTTTGGTTAAAAAATGTTTATTTGAA 300
QY	301 GTGCTTAAACAAAGAATATATTTCTGTGTATGTTAAATGGTTTGTGCCAACATGAATGG 360
DB	301 GTGCTTAAACAAAGAATATATTTCTGTGTATGTTAAATGGTTTGTGCCAACATGAATGG 360
QY	361 GGAAAAAGACCAAGCTGGCACTGCCATGFACTAAATTTGGAGGAAAGCACTTTAGTCAAGCT 420
DB	361 GGAAAAAGACCAAGCTGGCACTGCCATGFACTAAATTTGGAGGAAAGCACTTTAGTCAAGCT 420
QY	421 CAAGGGAAATGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGGTAACAGCC 480
DB	421 CAAGGGAAATGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGGTAACAGCC 480
QY	481 TGTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGACAAAT 540
DB	481 TGTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGACAAAT 540
QY	541 GAGTGGGTTACTCTACTTATTAAGCATAGCAAAACCAAAAAAGACTATACCAAGTGT 600
DB	541 GAGTGGGTTACTCTACTTATTAAGCATAGCAAAACCAAAAAAGACTATACCAAGTGT 600
QY	601 GTTCTTTTTGAAAACATGATGTTCTACTATTTTTTAACTTAAAAAGAAAATAAGCACTAGT 660
DB	601 GTTCTTTTTGAAAACATGATGTTCTACTATTTTTTAACTTAAAAAGAAAATAAGCACTAGT 660
QY	661 CCACCAAGAGACGGAGGCTATTTTCTTAGCAGTGA CTCTGGCTGGAAAACTAACTTTTTTA 720
DB	661 CCACCAAGAGACGGAGGCTATTTTCTTAGCAGTGA CTCTGGCTGGAAAACTAACTTTTTTA 720
QY	721 AAAGAGGCGAGCGCCATCTAGTGACCAACTATACACTGATGACATCGGCCGCAAGACG 780
DB	721 AAAGAGGCGAGCGCCATCTAGTGACCAACTATACACTGATGACATCGGCCGCAAGACG 780
QY	781 GTTGAACACCACTAACCCTGCGCAGGAAA CTAAAGCGCGGAGAAATTCAAACTAAAAAA 840
DB	781 GTTGAACACCACTAACCCTGCGCAGGAAA CTAAAGCGCGGAGAAATTCAAACTAAAAAA 840
QY	841 GAAGTTTCTATTAAAACTACACTTAAAGAGCTGGTGCA TAAAGAGTAACTCTCACAGAG 900
DB	841 GAAGTTTCTATTAAAACTACACTTAAAGAGCTGGTGCA TAAAGAGTAACTCTCACAGAG 900

QY 1981 TTGAAGAAAGACTTACGAGCGCGCTGAACCTTGACTTAA 2019
DB 1981 TTGAAGAAAGACTTACGAGCGCGCTGAACCTTGACTTAA 2019

RESULT 2
AAD02797
ID AAD02797 standard; DNA; 2019 BP.
AC AAD02797;
XX 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX Parvovirus non-structure protein 1 (NS1) wild-type DNA.
XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoural disease; gene therapy; ds.
XX Parvovirus.

Key Location/Qualifiers
CDS 1..2019
/*cag= a
/product= "Parvovirus NS1 protein"

EP1077260-A1.
XX 21-FEB-2001.
XX 13-AUG-1999; 99EP-00115161.
XX 13-AUG-1999; 99EP-00115161.
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX Nueesch J, Rommelaere J;
XX WPI: 2001-212717/22.
XX P-PSDB; AAY72702.
XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX Disclosure; Fig 1; 4lpp; English.
XX The present sequence is a wild type DNA encoding parvovirus non-
CC structure protein 1 (NS1). The present invention relates to the variants
CC of the parvovirus non-structure protein (NS1) having a shifted
CC equilibrium between the DNA replication and transcription activities, and
CC the cytotoxicity activity. These variants are useful as toxins for
CC treating tumoural diseases. The variant DNAs are useful as vectors for
CC gene therapy. (Updated on 06-AUG-2003 to correct OS field.)
XX Sequence 2019 BP; 698 A; 389 C; 452 G; 480 T; 0 U; 0 Other;

Query Match 99.9%; Score 2017.4; DB 5; Length 2019;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 2019; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA 60
DB 1 ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA 60
QY 61 AGTAACGAGAGTGTCTCATTGTTTTTAAATGAAATGTCAACTGAATGGAAAA 120
DB 61 AGTAACGAGAGTGTCTCATTGTTTTTAAATGAAATGTCAACTGAATGGAAAA 120
QY 121 GATATCGGATGGAATAGTTACAAAAAGAGCTGAGGAGCAGCTGAAATCTTTACAA 180
DB 121 GATATCGGATGGAATAGTTACAAAAAGAGCTGAGGAGCAGCTGAAATCTTTACAA 180

QY 181 CGAGGAGCGGAAACTACTTGGGACCAAAAGCGAGCAATGGAATGGGAAACCAACAGTGGAT 240
DB 181 CGAGGAGCGGAAACTACTTGGGACCAAAAGCGAGCAATGGAATGGGAAACCAACAGTGGAT 240
QY 241 GAAATGACCAAAAGCAAGTATTCTTTTGGTTAAATGTTTGGTTAAATGTTTAAATGTTTAA 300
DB 241 GAAATGACCAAAAGCAAGTATTCTTTTGGTTAAATGTTTGGTTAAATGTTTAAATGTTTAA 300
QY 301 GTGCTTAAACAAAGAAATATATTTCTCGTGTGATGTTAAATGTTTGGTTAAATGTTTAA 360
DB 301 GTGCTTAAACAAAGAAATATATTTCTCGTGTGATGTTAAATGTTTGGTTAAATGTTTAA 360
QY 361 GGAAGAGCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAAGGACTTTAGTCAAGCT 420
DB 361 GGAAGAGCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAAGGACTTTAGTCAAGCT 420
QY 421 CAAGGGAATGTTGGAGAGGCAACTAAATGTTTCTGGAGCAGATGTTGTTGTTAAACAGCC 480
DB 421 CAAGGGAATGTTGGAGAGGCAACTAAATGTTTCTGGAGCAGATGTTGTTGTTAAACAGCC 480
QY 481 TGTAAATGTCAACTAAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGACAAT 540
DB 481 TGTAAATGTCAACTAAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAGACAAT 540
QY 541 GAGTGGTTACTCTTACTTATAGCATAAGCAAAACCAAAAGACTATATACCAAGTGT 600
DB 541 GAGTGGTTACTCTTACTTATAGCATAAGCAAAACCAAAAGACTATATACCAAGTGT 600
QY 601 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAGAAATTAAGCACTAGT 660
DB 601 GTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAGAAATTAAGCACTAGT 660
QY 661 CCACCAAGAGACGAGCTATTTTCTTAGCAGTGACTCTGCTGGGAAATCTAACTTTTAA 720
DB 661 CCACCAAGAGACGAGCTATTTTCTTAGCAGTGACTCTGCTGGGAAATCTAACTTTTAA 720
QY 721 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACATGATGACATGCGGCCAGAAACG 780
DB 721 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACATGATGACATGCGGCCAGAAACG 780
QY 781 GTTGAACCAACAGTAAACCACTGGCGAGGAACTAAGCGGCGAGATTCAACTAAATAA 840
DB 781 GTTGAACCAACAGTAAACCACTGGCGAGGAACTAAGCGGCGAGATTCAACTAAATAA 840
QY 841 GAAGTTTCTATTAAACTACACTTAAAGAGCTGGTGATTAAGAGTAACTCAACAGAG 900
DB 841 GAAGTTTCTATTAAACTACACTTAAAGAGCTGGTGATTAAGAGTAACTCAACAGAG 900
QY 901 GACTGGATGATGATGAGCGAGCAGTTACATTTGAAATGATGGCTCAACAGGTGGAGAA 960
DB 901 GACTGGATGATGATGAGCGAGCAGTTACATTTGAAATGATGGCTCAACAGGTGGAGAA 960
QY 961 AACTGCTGAAAAATAGCTAGAGATTTGTACACTAACTCTAGCAGAACCAAAACAGCA 1020
DB 961 AACTGCTGAAAAATAGCTAGAGATTTGTACACTAACTCTAGCAGAACCAAAACAGCA 1020
QY 1021 TTTGACTTAAATTTAGAAAAAGCTGAAACCAAGCAAACTAACCAACTTTTCACTGCCTGAC 1080
DB 1021 TTTGACTTAAATTTAGAAAAAGCTGAAACCAAGCAAACTAACCAACTTTTCACTGCCTGAC 1080
QY 1081 ACAAGAACTGCAAGATTTTGTCTTTTGTGGTGGAACTATGTTAAAGTTTGGCATGCT 1140
DB 1081 ACAAGAACTGCAAGATTTTGTCTTTTGTGGTGGAACTATGTTAAAGTTTGGCATGCT 1140
QY 1141 ATTTGCTGTGTTTTAAACAGACAAAGGAGGCAAAAGAAATGCTTTTATTTATGAGGACCA 1200
DB 1141 ATTTGCTGTGTTTTAAACAGACAAAGGAGGCAAAAGAAATGCTTTTATTTATGAGGACCA 1200
QY 1201 GCCAGACAGGCAAACTATTATTGCAAGCCATAGCACAAGCAGTGGCAATGTTGGT 1260
DB 1201 GCCAGACAGGCAAACTATTATTGCAAGCCATAGCACAAGCAGTGGCAATGTTGGT 1260

Qy 361 GGAAGACCAAGCGCTGGCACTGTCATTAATTCGAGGAAAGGACTTTAGTCAAGCT 420
Db 621 GGAAGACCAAGCGCTGGCACTGTCATTAATTCGAGGAAAGGACTTTAGTCAAGCT 680
Qy 421 CAAGGGAATTTGGTGGAGAGGCACTAAATGTTTACTTGGAGCAGATGTTGGTAAACAGCC 480
Db 681 CAAGGGAATTTGGTGGAGAGGCACTAAATGTTTACTTGGAGCAGATGTTGGTAAACAGCC 740
Qy 481 TGTAAATGTCAACTTAACAACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAGCAAT 540
Db 741 TGTAAATGTCAACTTAACAACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAAGCAAT 800
Qy 541 GAGTGGGTTACTTACTTACTTATTAAGCATTAAGCAACCAACCAACCAACCAACCAACCA 600
Db 801 GAGTGGGTTACTTACTTACTTATTAAGCATTAAGCAACCAACCAACCAACCAACCAACCA 860
Qy 601 GTTCTTTTGGAAACATGATTTGCTTACTTATTTTAACTAAAGAAAGAAATTAAGCACTAGT 660
Db 861 GTTCTTTTGGAAACATGATTTGCTTACTTATTTTAACTAAAGAAAGAAATTAAGCACTAGT 920
Qy 661 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA 720
Db 921 CCACCAAGAGACGAGGCTATTTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAA 980
Qy 721 AAAGAGCGGAGCGCCATCTAGTGAGCAAACTATACATGATGACATGCGGCCGCAAGACG 780
Db 981 AAAGAGCGGAGCGCCATCTAGTGAGCAAACTATACATGATGACATGCGGCCGCAAGACG 1040
Qy 781 GTTGAACCAACAGTAACCACTGGCAGGAACTAAGCGCGCGAGAAATTCAACTTAAAAA 840
Db 1041 GTTGAACCAACAGTAACCACTGGCAGGAACTAAGCGCGCGAGAAATTCAACTTAAAAA 1100
Qy 841 GAAGTTTCTATTAACCACTAAGAGCTGGTGATTAAGAGAGTAACTCACCAGAG 900
Db 1101 GAAGTTTCTATTAACCACTAAGAGCTGGTGATTAAGAGAGTAACTCACCAGAG 1160
Qy 901 GACTGGATGATGACAGCAGAGTAACTGATGAATGATGGCTCAACAGGTGGAGAA 960
Db 1161 GACTGGATGATGACAGCAGAGTAACTGATGAATGATGGCTCAACAGGTGGAGAA 1220
Qy 961 AACTGCTGAAATATCGCTAGAGATTTGTACACTACTAGCCAGAACCAACAGCA 1020
Db 1221 AACTGCTGAAATATCGCTAGAGATTTGTACACTACTAGCCAGAACCAACAGCA 1280
Qy 1021 TTTGACTTAATTTAGAAAAAGCTGAAACCAAGCAAACTAACCACTTTCACTGCCTGAC 1080
Db 1281 TTTGACTTAATTTAGAAAAAGCTGAAACCAAGCAAACTAACCACTTTCACTGCCTGAC 1340
Qy 1081 ACAAGAACCTGCGAAATTTTGTCTTTTCATGGCTGGAACCTATGTTAAAGTTTGCATGCT 1140
Db 1341 ACAAGAACCTGCGAAATTTTGTCTTTTCATGGCTGGAACCTATGTTAAAGTTTGCATGCT 1400
Qy 1141 ATTTGCTGTTTAAACAGACAGGAGGCAAGAAATACTGTTTATTTCTAGGACCA 1200
Db 1401 ATTTGCTGTTTAAACAGACAGGAGGCAAGAAATACTGTTTATTTCTAGGACCA 1460
Qy 1201 GCCAGCAGCAGCAATCTATTATTCACAAGCCATAGCAGAGTGGCAATGTTGGT 1260
Db 1461 GCCAGCAGCAGCAATCTATTATTCACAAGCCATAGCAGAGTGGCAATGTTGGT 1520
Qy 1261 TGCTATAATGAGCCAAATGAACTTTTCCATTTTAACTGATGATCCCAACAGAACTTCATT 1320
Db 1521 TGCTATAATGAGCCAAATGAACTTTTCCATTTTAACTGATGATCCCAACAGAACTTCATT 1580
Qy 1321 TGGGTAGAAGAGCTGTAATCTTTGGACAGCAAGTAAACCAAGTTTAAAGCCATTGCTCT 1380
Db 1581 TGGGTAGAAGAGCTGTAATCTTTGGACAGCAAGTAAACCAAGTTTAAAGCCATTGCTCT 1640
Qy 1381 GGTCAAGCTATTGCACTTGTATCAAAAAAGGAAAGGAGCAAGAAACAGATTGAACCAACCA 1440
Db 1641 GGTCAAGCTATTGCACTTGTATCAAAAAAGGAAAGGAGCAAGAAACAGATTGAACCAACCA 1700
Qy 1441 GTCATCATGACCAAAATGAGAACTTATCAGTGGTTCAGAAATAGGCTCGGAAGAAAGCA 1500

Db 1701 GTCATCATGACCAAAATGAGAACTTACAGTGGTTCAGAAATAGGCTCGGAAGAAAGCA 1760
Qy 1501 GAACACACTCAACCAATTCAGAGACAGAAATGCTTAAACATTCATCTTAAACACATACCTTGCT 1560
Db 1761 GAACACACTCAACCAATTCAGAGACAGAAATGCTTAAACATTCATCTTAAACACATACCTTGCT 1820
Qy 1561 GGTGACTTTGGTTGGTTGTCACAAAATGAATGAGCCCAATGATTTGTGCTTGGTTGGTAAAG 1620
Db 1821 GGTGACTTTGGTTGGTTGTCACAAAATGAATGAGCCCAATGATTTGTGCTTGGTTGGTAAAG 1880
Qy 1621 AATGGTTACCAATCTACCATGCGCAAGCTACTGTGCTAAATGAGGCAAGTTTCTGATGG 1680
Db 1881 AATGGTTACCAATCTACCATGCGCAAGCTACTGTGCTAAATGAGGCAAGTTTCTGATGG 1940
Qy 1681 TCAGAAAACTGGCGGAGCCAAAGGTGCCAACTCCTATAAATTTACTAGGTTTGGCAGCG 1740
Db 1941 TCAGAAAACTGGCGGAGCCAAAGGTGCCAACTCCTATAAATTTACTAGGTTTGGCAGCG 2000
Qy 1741 TCACCAATTCACGACACCGAAAGTACGCCCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800
Db 2001 TCACCAATTCACGACACCGAAAGTACGCCCTCTCAGCCAGAACTATGCACTAACTCCACTT 2060
Qy 1801 GCATCGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACCAAAATCTCCTGTTGCG 1860
Db 2061 GCATCGATCTCGAGGACCTGGCTTTAGAGCCTTGGAGCACCAAAATCTCCTGTTGCG 2120
Qy 1861 GGCACTGCAAGAACCCAGAACTCTGGGAAAGCTGGTTTCCAAAGCTGCCAAGATGGTCAA 1920
Db 2121 GGCACTGCAAGAACCCAGAACTCTGGGAAAGCTGGTTTCCAAAGCTGCCAAGATGGTCAA 2180
Qy 1921 CTGAGCCCAACTTTGGTTCAGAGATCGAGGAGATTTGAGAGCGTCTTCGTTGCGGAACCG 1980
Db 2181 CTGAGCCCAACTTTGGTTCAGAGATCGAGGAGATTTGAGAGCGTCTTCGTTGCGGAACCG 2240
Qy 1981 TTGAAGAAAGACTTCAGCGAGCGCGCTGAACTTGGACTAA 2019
Db 2241 TTGAAGAAAGACTTCAGCGAGCGCGCTGAACTTGGACTAA 2279

RESULT 4
AAD02801 standard; DNA; 2019 BP.
XX AAD02801;
AC AC
XX XX
DT 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX Parvovirus non-structure protein 1 (NS1) variant (T363A) DNA.
DE NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoural disease; gene therapy; mutant; mutein; variant; ds.
XX Parvovirus.
OS Synthetic.
Key Location/Qualifiers
CDS 1..2019
FT /*tag= a
FT /product= "NS1 variant (T363A) protein"
FT mutation replace(1187, A)
FT /*tag= b
XX EP1077260-A1.
PN 21-FEB-2001.
XX 13-AUG-1999; 99EP-00115161.
XX 13-AUG-1999; 99EP-00115161.
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX	Nueesch J, Rommelaere J;			
PI	WPI: 2001-212717/22.			
XX	DR P-PSDB; AAY72706.			
XX	Novel parvovirus non-structure protein variant, useful for treating			
PT	tumoral diseases, has a shifted equilibrium between DNA replication and			
PT	transcription activities, and cytotoxic activity.			
XX	Claim 7; Page 16-19; 41pp; English.			
XX	The present sequence is a DNA encoding parvovirus non-structure protein 1			
CC	(NS1) variant (T363A). The invention relates to the variants of the			
CC	parvovirus non-structure protein (NS1) having a shifted equilibrium			
CC	between the DNA replication and transcription activities, and the			
CC	cytotoxicity activity. These variants are useful as toxins for treating			
CC	tumoral diseases. The variant DNAs are useful as vectors for gene			
CC	therapy. (Updated on 06-AUG-2003 to correct OS field.)			
XX				
SQ	Sequence 2019 BP; 697 A; 389 C; 453 G; 480 T; 0 U; 0 Other;			
Query Match 99.88; Score 2015.8; DB 5; Length 2019;				
Best Local Similarity 99.98; Pred. No. 0;				
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;				
Qy	1	ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA	60	
Dd	1	ATGGCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCAACTGGTTAAAGGAAAA	60	
Qy	61	AGTAACGAGGAAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGGAAAA	120	
Dd	61	AGTAACGAGGAAGTGTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGGAAAA	120	
Qy	121	GATATCGGATGGATAGTCTACAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTACAA	180	
Dd	121	GATATCGGATGGATAGTCTACAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTACAA	180	
Qy	181	CGAGGAGCGGAAACTACTTGGGACCAAGGAGGACATGGAATGGGAAACCAACAGTGGAT	240	
Dd	181	CGAGGAGCGGAAACTACTTGGGACCAAGGAGGACATGGAATGGGAAACCAACAGTGGAT	240	
Qy	241	GAAATGACCAAAAAGCAAGTATTCATTTTGTATTTTGTGTTTAAAAATGTTTATTGAA	300	
Dd	241	GAAATGACCAAAAAGCAAGTATTCATTTTGTATTTTGTGTTTAAAAATGTTTATTGAA	300	
Qy	301	GTGCTTTAAACAAGATATATTTCTGCTGATGTTAAATTTGTTTGCAACATGAAATGG	360	
Dd	301	GTGCTTTAAACAAGATATATTTCTGCTGATGTTAAATTTGTTTGCAACATGAAATGG	360	
Qy	361	GGAAAAACAAGCTGGCACTGCCATGTACTAATTCGAGGAAAGGACTTTAGTCAAGCT	420	
Dd	361	GGAAAAACAAGCTGGCACTGCCATGTACTAATTCGAGGAAAGGACTTTAGTCAAGCT	420	
Qy	421	CAAGGGAATGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGGTAAAGCC	480	
Dd	421	CAAGGGAATGGTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGGTTGGTAAAGCC	480	
Qy	481	TGTAATGTGCAACTAAACAAGCTGGAAGAAATTAACCTAAGAGAAATAGCAGAGACAAT	540	
Dd	481	TGTAATGTGCAACTAAACAAGCTGGAAGAAATTAACCTAAGAGAAATAGCAGAGACAAT	540	
Qy	541	GAGTGGGTTACTCTACTTACTTATAAGCATAAGCAAAACCAAAAGACTATACCAAGTGT	600	
Dd	541	GAGTGGGTTACTCTACTTACTTATAAGCATAAGCAAAACCAAAAGACTATACCAAGTGT	600	
Qy	601	GTTCTTTTGGAAACATGATTTGCTTACTATTTTAACTAAAGAAATAAGCACTAGT	660	
Dd	601	GTTCTTTTGGAAACATGATTTGCTTACTATTTTAACTAAAGAAATAAGCACTAGT	660	
Qy	661	CCACCAAGAGACGAGGCTATTTTCTTAGCAGTCACTCTGGCTGAAAACTAACTTTTAA	720	
Dd	661	CCACCAAGAGACGAGGCTATTTTCTTAGCAGTCACTCTGGCTGAAAACTAACTTTTAA	720	

QY 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGGAGCACACCAATACTCTCTTTGG 1860
Db 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGGAGCACACCAATACTCTCTTTGG 1860
QY 1861 GGCACTGCAGAAACCCAGAACACTCTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920
Db 1861 GGCACTGCAGAAACCCAGAACACTCTGGGGAAGCTGGTTCCAAAGCCTGCCAAGATGGTCAA 1920
QY 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGACGGTCTTCGGTGGGGAACCG 1980
Db 1921 CTGAGCCCAACTTGGTCAGAGATCGAGGAGGATTTGAGACGGTCTTCGGTGGGGAACCG 1980
QY 1981 TTGAAGAAAGACTTCAGCGAGCCGCTGAACCTTGGACTAA 2019
Db 1981 TTGAAGAAAGACTTCAGCGAGCCGCTGAACCTTGGACTAA 2019

RESULT 5
AAD02803
ID AAD02803 standard; DNA; 2019 BP.
AC AAD02803;
XX
XX
DT 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX
XX Parvovirus non-structure protein 1 (NS1) variant (T394A) DNA.
XX
XX NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoral disease; gene therapy; mutant; mutein; variant; ds.
XX
XX Parvovirus.
OS Synthetic.
OS
FH Key Location/Qualifiers
FT CDS l..2019
FT /*tag= a
FT /product= "NS1 variant (T394A) protein"
FT mutation replace(l180, A)
FT /*tag= b
XX
XX EP1077260-A1.
XX
XX 21-FEB-2001.
XX
XX 13-AUG-1999; 99EP-00115161.
XX
XX 13-AUG-1999; 99EP-00115161.
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
XX Nueesch J, Rommelaere J;
XX
XX WPI; 2001-212717/22.
DR P-PSDB; AAY72708.
XX
XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX
XX Claim 7; Page 22-24; 41pp; English.
XX
XX The present sequence is a DNA encoding parvovirus non-structure protein 1
CC (NS1) variant (T394A). The invention relates to the variants of the
CC parvovirus non-structure protein (NS1) having a shifted equilibrium
CC between the DNA replication and transcription activities, and the
CC cytotoxicity activity. These variants are useful as toxins for treating
CC tumoral diseases. The variant DNAs are useful as vectors for gene
CC therapy. (Updated on 06-AUG-2003 to correct OS field.)
XX
XX
XX Sequence 2019 BP; 697 A; 389 C; 453 G; 480 T; 0 U; 0 Other;
SQ Query Match 99.8%; Score 2015.8; DB 5; Length 2019;

Best Local Similarity 99.9%; Pred. No. 0;
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTTCGGAGCAACCAACTGGTTTAAAGGAAAA 60
Db 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTTCGGAGCAACCAACTGGTTTAAAGGAAAA 60
QY 61 AGTAAACCGAGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGGAATA 120
Db 61 AGTAAACCGAGAGTGTCTCATTTGTTTAAAAATGAAATGTTCAACTGAATGGAATA 120
QY 121 GATATCGGATGGAAATAGTTTACAAAAAGAGCTCGCAGAGGACGAGCTGAAATCTTTACAA 180
Db 121 GATATCGGATGGAAATAGTTTACAAAAAGAGCTCGCAGAGGACGAGCTGAAATCTTTACAA 180
QY 181 CGAGGAGCGGAAACTACTTTGGGACCAAGCGAGGACATGGAATGGGAAACCAACAGTGGAT 240
Db 181 CGAGGAGCGGAAACTACTTTGGGACCAAGCGAGGACATGGAATGGGAAACCAACAGTGGAT 240
QY 241 GAAATGACCAAAAAAGCAAGTATTTTTCATTTTGGTTTAAATGTTTATTTGAA 300
Db 241 GAAATGACCAAAAAAGCAAGTATTTTTCATTTTGGTTTAAATGTTTATTTGAA 300
QY 301 GTGCTTAAACACAAAGAAATATATTTCTGTTGATGTTTAAATGTTTGGCAACATGAATGG 360
Db 301 GTGCTTAAACACAAAGAAATATATTTCTGTTGATGTTTAAATGTTTGGCAACATGAATGG 360
QY 361 GGAAGAGACCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAAGACTTTTAGTCAAGCT 420
Db 361 GGAAGAGACCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAAGACTTTTAGTCAAGCT 420
QY 421 CAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGTTGGTAAACAGCC 480
Db 421 CAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGTTGGTAAACAGCC 480
QY 481 TGTAAATGTCGAACCTACACACGCTGAAAGAAATTAAGTAAGCAATAGCAGAGCAAT 540
Db 481 TGTAAATGTCGAACCTACACACGCTGAAAGAAATTAAGTAAGCAATAGCAGAGCAAT 540
QY 541 GAGTGGTGTACTCTACTTACTTATAAGCATATAAGCAACCAAAAAAGACTATACCAAGTGT 600
Db 541 GAGTGGTGTACTCTACTTACTTATAAGCATATAAGCAACCAAAAAAGACTATACCAAGTGT 600
QY 601 GTTCTTTTGGAAACATGATTTGCTTACTATTTTAACTAAAAAGAAATTAAGCACTAGT 660
Db 601 GTTCTTTTGGAAACATGATTTGCTTACTATTTTAACTAAAAAGAAATTAAGCACTAGT 660
QY 661 CCACCAAGAGACGGAGGCTATTTTCTTAGCAGTACTCTGGCTGGAAACCTACTTTTAA 720
Db 661 CCACCAAGAGACGGAGGCTATTTTCTTAGCAGTACTCTGGCTGGAAACCTACTTTTAA 720
QY 721 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATCGGCCAGAAACG 780
Db 721 AAAGAGGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATCGGCCAGAAACG 780
QY 781 GTTGAACCAACAGTAAACCACTGCGCAGGAACTAAGCGCGCAGAAATTCAAACTAAAAA 840
Db 781 GTTGAACCAACAGTAAACCACTGCGCAGGAACTAAGCGCGCAGAAATTCAAACTAAAAA 840
QY 841 GAAGTTTCTATTAAAACTACCTTAAGAGCTGGTGATAAAAAGAGTAACCTCACCAGAG 900
Db 841 GAAGTTTCTATTAAAACTACCTTAAGAGCTGGTGATAAAAAGAGTAACCTCACCAGAG 900
QY 901 GACTGATGATGATGAGCCAGACAGTTACATTGAAATGATGCTCAACCCAGGTGGAGAA 960
Db 901 GACTGATGATGATGAGCCAGACAGTTACATTGAAATGATGCTCAACCCAGGTGGAGAA 960
QY 961 AACCTGCTGAAAAATACGCTTAGAGATTTGTACAACTACTAGCCAGAACCAAAACAGCA 1020
Db 961 AACCTGCTGAAAAATACGCTTAGAGATTTGTACAACTACTAGCCAGAACCAAAACAGCA 1020
QY 1021 TTTGACTTAAATTTAGAAAAAGCTGAAACCCAGCAAACTAACCACTTTTCACTGCTTGAC 1080
Db 1021 TTTGACTTAAATTTAGAAAAAGCTGAAACCCAGCAAACTAACCACTTTTCACTGCTTGAC 1080

Db 1021 TTTGACTTAATTTTAAAGAAAGCTGAAACACGAGCAAACTAACCAACTTTTTCACCTGCCTGAC 1080
Qy 1081 ACAGAAACCTGCAGAAATTTTGTCTTTTCATGGCTGGAACTATGTTAAAGTTTGGCAATGCT 1140
Db 1081 ACAGAAACCTGCAGAAATTTTGTCTTTTCATGGCTGGAACTATGTTAAAGTTTGGCAATGCT 1140
Qy 1141 ATTGCTGTGTTTAAACACAGAGGAGGCAAAAGAAATCTGTTTATTTATTCATGAGACCA 1200
Db 1141 ATTGCTGTGTTTAAACACAGAGGAGGCAAAAGAAATCTGTTTATTTATTCATGAGACCA 1200
Qy 1201 GCCAGCAGCGCAATCTATTTATTCACAGCCATAGCACAGCAGTTGGCAATGTTGCT 1260
Db 1201 GCCAGCAGCGCAATCTATTTATTCACAGCCATAGCACAGCAGTTGGCAATGTTGCT 1260
Qy 1261 TGCTATAATGAGCCATGTAACTTTTCCATTTTAACTGTTACCAAGCAAGCACTTGATT 1320
Db 1261 TGCTATAATGAGCCATGTAACTTTTCCATTTTAACTGTTACCAAGCAAGCACTTGATT 1320
Qy 1321 TGGGTAGAAGAGCTGTTAACTTTTGGACAGCAAGTAAACAGTTTAAAGCCATTGCTCT 1380
Db 1321 TGGGTAGAAGAGCTGTTAACTTTTGGACAGCAAGTAAACAGTTTAAAGCCATTGCTCT 1380
Qy 1381 GGTCAAGCTATTGCAATGTATCAAAAGGAAAGGCAAGCAAGATTTGAACCAACACCA 1440
Db 1381 GGTCAAGCTATTGCAATGTATCAAAAGGAAAGGCAAGCAAGATTTGAACCAACACCA 1440
Qy 1441 GTCATATGACCAAAATGAGNACATTACAGTGTGTCAGATAGGCTGCGAAGAAAGACCA 1500
Db 1441 GTCATATGACCAAAATGAGNACATTACAGTGTGTCAGATAGGCTGCGAAGAAAGACCA 1500
Qy 1501 GAACACACTCAACCAATCAGAGACAGAACTGTTAAACATTCATCTAACACATACCTTGCT 1560
Db 1501 GAACACACTCAACCAATCAGAGACAGAACTGTTAAACATTCATCTAACACATACCTTGCT 1560
Qy 1561 GGTGACTTTGGTTGTCAGAAATGAATGCGCCATGATTTGCTGTTGGTAAAG 1620
Db 1561 GGTGACTTTGGTTGTCAGAAATGAATGCGCCATGATTTGCTGTTGGTAAAG 1620
Qy 1621 AATGTTTACCAATCTACATGGCAAGCTACTGTCTAAATGGGCAAAAGTTCTCTGATTGG 1680
Db 1621 AATGTTTACCAATCTACATGGCAAGCTACTGTCTAAATGGGCAAAAGTTCTCTGATTGG 1680
Qy 1681 TCAGAAACCTGGGCGAGCAAGGTCGCAACTCTCTATAAATTTACTAGTTTGGCAGCG 1740
Db 1681 TCAGAAACCTGGGCGAGCAAGGTCGCAACTCTCTATAAATTTACTAGTTTGGCAGCG 1740
Qy 1741 TCACCAATTCAGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800
Db 1741 TCACCAATTCAGACACCGAAAGTACGCTCTCAGCCAGAACTATGCACTAACTCCACTT 1800
Qy 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAATATCTCTGTTGCG 1860
Db 1801 GCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAATATCTCTGTTGCG 1860
Qy 1861 GGCACTCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTTGCCAAAGATGTCFAA 1920
Db 1861 GGCACTCAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCTTGCCAAAGATGTCFAA 1920
Qy 1921 CTGAGCCCACTTGGTCAGAGATCGAGGAGATTTGAGCGCTGCTTCGGTGGGAACCG 1980
Db 1921 CTGAGCCCACTTGGTCAGAGATCGAGGAGATTTGAGCGCTGCTTCGGTGGGAACCG 1980
Qy 1981 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019
Db 1981 TTGAAGAAAGACTTTCAGCGAGCGCTGAACTTGGACTAA 2019

RESULT 6

AAD02799

ID AAD02799 standard; DNA; 2019 BP.

XX

AC AAD02799;

XX

DT 06-AUG-2003 (revised)
DT 31-MAY-2001 (first entry)
XX Parvovirus non-structure protein 1 (NS1) variant (S283A) DNA.
DE NS1; non-structure protein 1; cytostatic; gene therapy; toxin; therapy;
KW tumoral disease; gene therapy; mutant; mutein; variant; ds.
XX Parvovirus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..2019
FT /tag= a
FT /product= "NS1 variant (S283A) protein"
FT mutation replace(847, A)
FT /tag= b
XX
PN EP1077260-A1.
XX
PD 21-FEB-2001.
XX
PF 13-AUG-1999; 99BP-00115161.
XX
PR 13-AUG-1999; 99BP-00115161.
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
XX Nueesch J, Rommelaere J;
XX
XX WPI; 2001-212717/22.
DR P-PSDB; AAY72704.
XX
XX Novel parvovirus non-structure protein variant, useful for treating
PT tumoral diseases, has a shifted equilibrium between DNA replication and
PT transcription activities, and cytotoxic activity.
XX
XX Claim 7; Page 11-14; 41pp; English.
XX
XX The present sequence is a DNA encoding parvovirus non-structure 1 protein
CC (NS1) variant (S283A). The invention relates to the variants of the
CC parvovirus non-structure protein (NS1) having a shifted equilibrium
CC between the DNA replication and transcription activities, and the
CC cytotoxicity activity. These variants are useful as toxins for treating
CC tumoral diseases. The variant DNAs are useful as vectors for gene
CC therapy. (Updated on 06-AUG-2003 to correct OS field.)
XX
XX Sequence 2019 BP; 698 A; 389 C; 453 G; 479 T; 0 U; 0 Other;
SQ
Query Match 99.8%; Score 2015.8; DB 5; Length 2019;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2017; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGGCTCGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGTTAAAGGAAAA 60
Db 1 ATGGCTCGAAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGTTAAAGGAAAA 60
Qy 61 AGTAACACGAGAGTGTCTCATTTGTTTAAAAATGAAAAATGTTCAACTGAATGAAAA 120
Db 61 AGTAACACGAGAGTGTCTCATTTGTTTAAAAATGAAAAATGTTCAACTGAATGAAAA 120
Qy 121 GATATCGGATGGAATAGTTACAAAAAGAGCTGCAGGAGACGAGCTGAAATCTTTACAA 180
Db 121 GATATCGGATGGAATAGTTACAAAAAGAGCTGCAGGAGACGAGCTGAAATCTTTACAA 180
Qy 181 CGAGGAGCGGAACACTTGGGACCAAGCGAGGACATGGAATGGGAACCAAGTGGAT 240
Db 181 CGAGGAGCGGAACACTTGGGACCAAGCGAGGACATGGAATGGGAACCAAGTGGAT 240
Qy 241 GAAATGACCAAAAGCAAGTATTTCATTTTGGTTTAAAAAATGTTTATTGAA 300
Db 241 GAAATGACCAAAAGCAAGTATTTCATTTTGGTTTAAAAAATGTTTATTGAA 300

Qy	301	GTGCTTAAACAAAGAATATATATTCCTCGGTGATGTATATTCCTGGTTGTGCAACATGAATGG	360
Db	301	GTGCTTAAACAAAGAATATATATTCCTCGGTGATGTATATTCCTGGTTGTGCAACATGAATGG	360
Qy	361	GGAAAAAGCAAGGCTGCACCTGCCATGTACTAATTTGGAGGAAAGGACTTTTAGTCAAGCT	420
Db	361	GGAAAAAGCAAGGCTGCACCTGCCATGTACTAATTTGGAGGAAAGGACTTTTAGTCAAGCT	420
Qy	421	CAAGGGAAATGGTTGGAGAAAGCAACTAAATGTTTACTTGGAGCAGATGGTTGGTAAACAGCC	480
Db	421	CAAGGGAAATGGTTGGAGAAAGCAACTAAATGTTTACTTGGAGCAGATGGTTGGTAAACAGCC	480
Qy	481	TGTTAATGTGCAACTAAACACAGCTGAAAGAAATTTAAACTAAGAGAAATAGCGAAGACAAT	540
Db	481	TGTTAATGTGCAACTAAACACAGCTGAAAGAAATTTAAACTAAGAGAAATAGCGAAGACAAT	540
Qy	541	GAGTGGGTTACTCTTACTTTACTTTAAGCATAAGCAAAACCAAAAAGAGACTATACCAAGTGT	600
Db	541	GAGTGGGTTACTCTTACTTTACTTTAAGCATAAGCAAAACCAAAAAGAGACTATACCAAGTGT	600
Qy	601	GTTCTTTTTTGGAAAACATGATTGCTTACTATTTTAACTAAAAAGAAATAAGCACTAGT	660
Db	601	GTTCTTTTTTGGAAAACATGATTGCTTACTATTTTAACTAAAAAGAAATAAGCACTAGT	660
Qy	661	CCACCAAGAGACGGAGGCTATTTCTTTAGCAGTGACTCTGGCTGGAAAACTTAACCTTTTTA	720
Db	661	CCACCAAGAGACGGAGGCTATTTCTTTAGCAGTGACTCTGGCTGGAAAACTTAACCTTTTTA	720
Qy	721	AAAGAGCGGAGCGCCATCTAGTCAGACAAACTATACACTGATGATGATCGCGGCAGAAACG	780
Db	721	AAAGAGCGGAGCGCCATCTAGTCAGACAAACTATACACTGATGATGATCGCGGCAGAAACG	780
Qy	781	GTTGAAACCAACAGTAAACCACTGGCCAGGAAACTAAGCGCGGCAGAACTCAAACTAAAAA	840
Db	781	GTTGAAACCAACAGTAAACCACTGGCCAGGAAACTAAGCGCGGCAGAACTCAAACTAAAAA	840
Qy	841	GAACTTCTATTAATAACTACACTTTAAAGAGCTGGTGCATATAAGAGTAACTCACACAGAG	900
Db	841	GAACTTCTATTAATAACTACACTTTAAAGAGCTGGTGCATATAAGAGTAACTCACACAGAG	900
Qy	901	GACTGGATGATGTCGACGACAGATTACATTTCAAAATGATGGCTCAACAGGTGGAGAA	960
Db	901	GACTGGATGATGTCGACGACAGATTACATTTCAAAATGATGGCTCAACAGGTGGAGAA	960
Qy	961	AACTGCTGAAAAATAGCGTAGAGATTGTACATAACTCTAGCCAGAAACCAAAACAGCA	1020
Db	961	AACTGCTGAAAAATAGCGTAGAGATTGTACATAACTCTAGCCAGAAACCAAAACAGCA	1020
Qy	1021	TTTGACTTAATTTTGAAGAAAGCTGAAACACAGCAAACTAACCAACTTTTCACTGCCTGAC	1080
Db	1021	TTTGACTTAATTTTGAAGAAAGCTGAAACACAGCAAACTAACCAACTTTTCACTGCCTGAC	1080
Qy	1081	ACAGAACCTGACAGAAATTTTGTCTTTTCATGGCTGGAACTATGTTAAAGTTTGCCATGCT	1140
Db	1081	ACAGAACCTGACAGAAATTTTGTCTTTTCATGGCTGGAACTATGTTAAAGTTTGCCATGCT	1140
Qy	1141	ATTTGTCTGTTTTTAAACAGACAGGAGGCAAAAGAAATACTGTTTATTTTCAATGGACCA	1200
Db	1141	ATTTGTCTGTTTTTAAACAGACAGGAGGCAAAAGAAATACTGTTTATTTTCAATGGACCA	1200
Qy	1201	GCCAGCACAGGCAAACTCTATTTATGACAGCCCATAGCAACAAGCAGTTGGCAATTTGGT	1260
Db	1201	GCCAGCACAGGCAAACTCTATTTATGACAGCCCATAGCAACAAGCAGTTGGCAATTTGGT	1260
Qy	1261	TGCTTATTAATGCAGCCAATGTAACCTTTCCATTTTAATGACTGTATACCAACAGAACTTGATT	1320
Db	1261	TGCTTATTAATGCAGCCAATGTAACCTTTCCATTTTAATGACTGTATACCAACAGAACTTGATT	1320
Qy	1321	TGGGTAGAAAGAGCTGTTGACTTTGGACAGCAAGTAAACCAAGTTTAAAGCCATTTGCTCT	1380
Db	1321	TGGGTAGAAAGAGCTGTTGACTTTGGACAGCAAGTAAACCAAGTTTAAAGCCATTTGCTCT	1380
Qy	1381	GGTCAAGCTATTTGCGATTGATCAAAAAAGGAAAAAGGCGCAAAACAGATTTGAACCAACCA	1440

Db	1381	GGTCAAACTATTGTCGATTTGATCAAAAGGAAAGGCGAGCAACAGATTTGAACCAACCA	1440
Qy	1441	GTCAATCATGACCACAAATGAGAACATTTACAGTTGGTCAGAAATAGGCTCGGAAGAAAGACCA	1500
Db	1441	GTCAATCATGACCACAAATGAGAACATTTACAGTTGGTCAGAAATAGGCTCGGAAGAAAGACCA	1500
Qy	1501	GAACACACTCAACCAATCAGAGACAGAAATGCTTTAACTTCACTTAAACACATACCTTGCCT	1560
Db	1501	GAACACACTCAACCAATCAGAGACAGAAATGCTTTAACTTCACTTAAACACATACCTTGCCT	1560
Qy	1561	GGTGACTTTGGTTGGTTGTGACAAAATGAAATGGCCCATGATTTGTGCTCTGGTTGGTAAAG	1620
Db	1561	GGTGACTTTGGTTGGTTGTGACAAAATGAAATGGCCCATGATTTGTGCTCTGGTTGGTAAAG	1620
Qy	1621	AATGGTTACCAATCTACCATTGGCGAAGCTACTGTGCTTAAATTTGGGCGCAAGTTCTCTGATGG	1680
Db	1621	AATGGTTACCAATCTACCATTGGCGAAGCTACTGTGCTTAAATTTGGGCGCAAGTTCTCTGATGG	1680
Qy	1681	TCAGAAAATCTGGGCGGAGCCAAAGGTGCCAACTCCTATAAATTTACTTAGGTTGGGCACGC	1740
Db	1681	TCAGAAAATCTGGGCGGAGCCAAAGGTGCCAACTCCTATAAATTTACTTAGGTTGGGCACGC	1740
Qy	1741	TCACCATTCACGACACCGGAAAGTACGCCCTCTCGAGCCAGAACTATGCACTTAACTTCCAATT	1800
Db	1741	TCACCATTCACGACACCGGAAAGTACGCCCTCTCGAGCCAGAACTATGCACTTAACTTCCAATT	1800
Qy	1801	GCATCGGATCTCGAGGACCTGGCTTTTAGAGCCTTTGGAGCACACAAATACTCTCTGTTGCG	1860
Db	1801	GCATCGGATCTCGAGGACCTGGCTTTTAGAGCCTTTGGAGCACACAAATACTCTCTGTTGCG	1860
Qy	1861	GGCACTGCGAAGAACCCAGAAACA CTGGGGAGAGCTGGTTTCAAAGCCTGCCAAGATGTTCAA	1920
Db	1861	GGCACTGCGAAGAACCCAGAAACA CTGGGGAGAGCTGGTTTCAAAGCCTGCCAAGATGTTCAA	1920
Qy	1921	CTGAGCCCACTTTGGTCAGAGATTCGAGGAGGATTTGAGACCGTGCTTTCCGTCGGGACCG	1980
Db	1921	CTGAGCCCACTTTGGTCAGAGATTCGAGGAGGATTTGAGACCGTGCTTTCCGTCGGGACCG	1980
Qy	1981	TTGAAGAAAGACTTCAGCGAGCCGCTGAACTTCGACTAA	2019
Db	1981	TTGAAGAAAGACTTCAGCGAGCCGCTGAACTTCGACTAA	2019

RESULT 7
AAT15311
ID AAT15311 standard; DNA; 5049 BP.

```

RESULT 7
AAT15311
ID AAT15311 standard; DNA; 5049 BP.
XX
XX
AC AAT15311;
XX
XX
DT 14-OCT-1996 (first entry)
XX
XX
DE Non-attenuated canine parvovirus CPV-39 passage 5 DNA.
XX
XX
KW Parvovirus; dog; vaccine; CPV; myocardial disease; enteric disease; ds;
KW ss.
XX
XX
OS Canine parvovirus.
XX
XX
FN WO9614088-A1.
XX
XX
PD 17-MAY-1996.
XX
XX
XX 02-NOV-1995; 95WO-US014207.
XX
XX
PR 08-NOV-1994; 94US-00336345.
XX
XX
(CORR ) CORNELL RES FOUND INC.
XX
XX
PI Parrish CR, Gruenberg A, Carmichael LE;
XX
XX
WPI; 1996-251556/25.
XX
XX

```

PT Attenuated CPV strains contg. up to 4 mutation (s) relative to control
PT virus - useful as a veterinary vaccine against CPV disease in animals,
PX such as wild or domestic dogs.

PS Claim 1; Page 21-24; 42pp; English.

XX This viral DNA is isolated from a non- attenuated CPV. The DNA is
CC preferably derived from vB1440. The DNA is cloned into a vector which is
CC used to transfect a host cell. The vector used is preferably pGEM3Z or
CC pGEM5Z. The host cells to be transfected are selected from Norden
CC Laboratory feline kidney cells, mink lung cells, Madin-Darby canine
CC kidney cells or canine A72 cells

XX SQ Sequence 5049 bp; 1791 A; 812 C; 1029 G; 1417 T; 0 U; 0 Other;

Query Match 51.5%; Score 1040.4; DB 2; Length 5049;

Best Local Similarity 71.4%; Pred. No. 8.9e-270;

Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;

QY	1	ATGGCTGGAAATGCTTACTCTGATGAAGTTTGGGACCAACCACTGGTTAAAGGAAAAA	60
DB	273	ATGCTGGCAACGATATCTGGAAGTTATGGAGGAGTAAATGGTTAAGAAACAT	332
QY	61	AGTAACGAGGAGTGTCTCATTTGTTTTTAAATAAGAAATGTTCAACTGGAATGGAAAA	120
DB	333	GCAGAAATGAAGCATTTTCGTTGTTTTTAAATGTGACCAACGTCCAACTAAATGGAAAG	392
QY	121	GATATCGGATGGATAGTACMAAAGAGCTGCAGGAGCGAGCTGGAATCTTTACAA	180
DB	393	GATGTTGCTGGAACTACTATACCAACCAATTCMAAATGAAGCTAACATCTTTAAAT	452
QY	181	CGAGGAGCGGAAACTACTTGGGACCAAGC-----GAGGACATGGAATGGGAACCA	234
DB	453	AGAGGAGCAACAACAGCAATGGATCAACCGAAGAGAAATGGATCGGAATCGGAA	512
QY	235	GTGATGAATGACCAAAAAAGCAAGTATTCATTTTGTTGTTGTTTAAATAATGTTTA	294
DB	513	GTTGATAGTCTCGCCAAAAAGCAAGTACAACTTTTGATGCAATTAATTAATAAATGTC	572
QY	295	TTTGAAGTCTTAACACAGAGATATATTTCTGGTGCATGTTAATGGTTTGGCAACAT	354
DB	573	TTTGAAGTCTTTGTTCTTAAATATATAGAACCAATGAATGTTGTTGTTTATTCACAT	632
QY	355	GAATGGGAAAAAGACCAAGCTGGCACTGGCATGTACTTAATTTGAGGAAAGCACTT	414
DB	633	GAATGGGAAAAAGATCAAGCTGGCATGTGATGTTTCTTCACTAGTAAGACTTACAA	692
QY	415	CAAGCTCAAGGAAATGGTGGAGAGGCAACTAAATGTTTACTTGGAGCAGATGGTGGTA	474
DB	693	CAAGCAACTGGTAAATGGCTACGACAGACAAATGAATATGTTTGGAGTAGATGGTGG	752
QY	475	ACAGCTGTAATGTGCACTAACACACAGCTGAAGAAATTAACCTAAGAGAAATAGCAGAA	534
DB	753	ACTCTTTGTTGGTAAACTTAAACACCACTGAAAGATTAAGCTCAGAGAAATTTGCA	812
QY	535	GACAACTGAGTGGGTACTCTACTTACTTATAAGCATAAAGCAACCAAAAGCACTATAC	594
DB	813	GATAGTGAATGGGTGACTATATTAACATACAGACATGAAGCAACAAAGAAAGACTAT	872
QY	595	AAGTGTGTTCTTTTGGAAACATGATGCTTACTATTTTAACTTAAAGAAATAAGC	654
DB	873	AAAATGTTTCAATTTGGAAATATGATAGCATATTTACTTTTAAACAAAGAAAAATGTC	932
QY	655	ACTAGTCCACCAAGAGAGCGGCTATTTCTTAGCAGTGAATCTGCTGCGGAAACTTAAC	714
DB	933	CACAT---GACAAAGAAAGTGGCTATTTTAAAGTACTGATCTGTTGGAATTTAAC	989
QY	715	TTTTTAAAGAGCGGAGCGCACTCTAGTGAGCAAACTATACACTGATGACATGCGGCCA	774
DB	990	TTTATGAAGTATCAAGACAGACAAATTTGTGAGCACACTTTACACTGAAACAAATGA	1049
QY	775	GAACCGTTGMAACCAAGTAAACACTGCGGAGGAACTAAGCGCGGCAGAAATTCAACT	834

DB	1050	GAACCGTTGAAACCAACAGACGACGACGACAGAAACAAAGCGCGGAGAAATTCAAACT	1109
QY	835	AAAAAGAGAGTTTCTATTAAAACTACACTTAAAGAGCTGTCATATAAAGAGTAACCTCA	894
DB	1110	AAAAAGAGTGTCAATCAATGTACTTTGCGGAGCTTGGTTAGTAAAGAGTAAACATCA	1169
QY	895	CCAGAGGACTGGATGATGATGCAGCCAGACAGTTACATTGAAATGATGGCTCAACGAGT	954
DB	1170	CCTGAAGACTGGATGATGTTTACAAACAGATAGTTATATTGAAATGATGGCACAACCA	1229
QY	955	GGAGAAAACTGCTGAAAAATAGCTAGAGATTGTACACTACTCTAGCCAGAACCAAA	1014
DB	1230	GGTGAAAAATCTTTTAAAAAATACACTTGAATTTGTACTTTGACTTTAGCAAGACAAA	1289
QY	1015	ACAGCATTTGACTTAAATTTTAGAAAAAGCTGAAACCGACAACTAACCAACTTTTCACTG	1074
DB	1290	ACAGCATTTGAAATTAATTAATCTTGAAGGAGAGATATATCTAACTAACTTAATTTG	1349
QY	1075	CCTGACACAAGAACCTGCAGAAATTTTGTCTTTTTCATGGCTGGAACTATGTTAAAGTTGC	1134
DB	1350	GCAAAATTTAGAACATGTCAAAATTTTGAATGCGAGGATGGAATTTGATTTAAAGTTGT	1409
QY	1135	CATGCTATTTGCTGCTGTTTAAACAGACAAAGGAGGCAAAAGAAATACTGTTTATTTTCAT	1194
DB	1410	CACGCTATAGCATGTGTTTTTAAATAGCAAGGTGGTAAAGAAATACAGTTCTTTTTCAT	1469
QY	1195	GGACGAGCAGCAGCGCAAACTCTATTATTGCAAGCCATAGCAAGCAGTGTGGCAAT	1254
DB	1470	GGACGAGCAAGTACAGGAAATCTATCTTCTCAAGCCATAGCAAGCTGTGGGTAAAT	1529
QY	1255	GTTGTTGCTTAAATCAGCCAAATGTAATTTTCAATTTTAACTGTGTACCAACAAGAAC	1314
DB	1530	GTTGTTGTTTAAATCAGCAAAATGTAATTTTCCATTTTAACTGTGTACCAATAAAAT	1589
QY	1315	TTGATTTGGGTAGAAAGCTGTAATTTTGGACAGCAAGTAAACAGTTTAAAGCCATT	1374
DB	1590	TTAAATTTGGAATGGAAGCTGGTAACTTTGGTCAACAAAGTTAATCAATTTAAAGCAATC	1649
QY	1375	TGCTCTGGTCAAGCTATTCGCTTGAATCAAAAGGAAAGGAGCAAGCAAGATTTGAACCA	1434
DB	1650	TGTTCTGGACAAACAAATTAAGATTGATCAAAAGGTAAAGGAAGTAAAGCAAAATTTGAACCA	1709
QY	1435	ACAACGATCATATGACCAACAATGAGAACATTTACAGTGGTCAAGATAGCTCGGAGAA	1494
DB	1710	ACTCCAGTAAATATGACAACTAATGAAATATAACAAATTTGTGAGAAATTTGGATGTGAAGAA	1769
QY	1495	AGACGAGAACACACTCAACCAATCAGACAGCAAGTCTTAAACATTTCTATCAACATACC	1554
DB	1770	AGACCTGAAACATACACAACCAATAGAGACAGAAATGTTGAACATTAAGTTAGTATGTAAG	1829
QY	1555	TTGCTCTGGTGAATTTGGTTGGTGAACAAAATGAAATGGCCCATGATTTGTCTGGTTG	1614
DB	1830	CTTCCAGGAGACTTTGGTTGGTTGATAAAGAAATGGCTTTAATATGTGATGGTTA	1889
QY	1615	GTAAGAATGGTTTCAATCTTACATGGCAAGCTACTGTGTCTAAATTTGGGGCAAAAGTTCTCT	1674
DB	1890	GTTAAACATGGTTTGAATCAACCATGGCTTAACATATACACATCATTTGGGGAAAAAGTACCA	1949
QY	1675	GATTTGTCAGAAACTGGGGGAGCCAAAGGTGCGCACTCTCTATAAATTTACTAGTTTCG	1734
DB	1950	GAATGGATGAAAACTGGGGGAGCCCTAAATAACAAGAGGTATAAATCCACCGGTTCG	2009
QY	1735	GCAGCTCACCACTTACGACACCGAAAAAGTACGCTCTCAGCCAGCAAACTATGCACTAACT	1794
DB	2010	AAAGCT---TAGAGACACAGCGGCAAGCAATCTCTCAGAGTCAAGACCAAGTTCTAACT	2066
QY	1795	CACTTTGCAATCGGATCTGAGGACCTGGCTTTAGAGCTTTGAGGACACCAAAATCTCTCT	1854
DB	2067	CCTCTGACTCCGACCGTAGTGGACCTTGCACTGGAACCGGTGGAGTACTCCAGATAGCCT	2126
QY	1855	GTTTGGCGGCACTGCAGAAACCCAGAACACTGCGGGAAGCTGGTTCCAAAGCCTGCCAAGAT	1914
DB	2127	ATTGCAAGAACTGCAAAATCAACAACTTGGGCTTACTCA-----CAAGAC	2180

Db 1350 GCAATTCCTAGAACATGTCAAATTTTGTAGAAATGCACGGATGGAATTTGGAATTAAGTTTGT 1409
QY 1135 CATGCTATTTCGTGTTTAAACACAGCAGGAGGCAAAAGAAATACTGTTTATTTTCAT 1194
Db 1410 CAGCCTATAGCATGTGTTTAAATAGACAGGTGGTAAAGAAATACAGTCTCTTTTCAT 1469
QY 1195 GGACGACCCAGCAGCAGCAAACTATATTTGCAAGCCCATAGCAAGCAGTTGGCAAT 1254
Db 1470 GGACGACGAGTACAGGAAATCTATCTGCTCAAGCCATAGCAAGCTGTGGGTAAT 1529
QY 1255 GTTGGTGTCTATATGCGCCATGTAACTTTCCATTTTAACTGACGTACCAACAGAAC 1314
Db 1530 GTTGGTGTCTATATGCGCAATGTAAATTTTCCATTTTAACTGACGTACCAATAAAAT 1589
QY 1315 TTCAATTTGGGTAGAGAGCTGTAACTTTGGGACAGCAAGTAAACAGTTTAAAGCCATT 1374
Db 1590 TTAATTTGGATTGAGAGCTGTAACTTTGGTCAACAGTTAATCAATTTAAAGCAATC 1649
QY 1375 TGCTCTGGTCAAGCTATTCGATTTGATCAAAAGGAAAGGCAAGCAAGATTTGAACCA 1434
Db 1650 TGTCTGGACAAACAATTAGAAATGATCAAAAGGTAAGGAAGTAAGCAAAATTTGAACCA 1709
QY 1435 ACACGAGTATCATGACCAACAATGAGAACATTTACAGTGTGAGATAGGCTCGGAAGAA 1494
Db 1710 ACTCCAGTAATTTATGCAACTAATGAAATATTAACAATTTGAGAAATTTGATGTGAAGAA 1769
QY 1495 AGACCAAGACACACTCAACCAATCAGACAGCAAGTCTTAACTTCACTTAACACATACC 1554
Db 1770 AGACCTGAACATACACACCAATTAAGAGACAGAAATGTTGACATTTAGTAGTATGTAAG 1829
QY 1555 TTGCTCTGGTCACTTTGGTTGGTTGACAAAAATGAATGGCCCAATGTTGTTGGTTG 1614
Db 1830 CTTCCAGGACACTTTGGTTGGTTGATAAAGAAAGTGGCTTTAATATGTCGATGGTTA 1889
QY 1615 GTAAGAGATGTTACCAATCTACATGCGCAAGTACTGTCTAAATGGGCAAGTTTCTCT 1674
Db 1890 GTTAAACATGGTTTGAATCAACCATGGCTTAACATACATCATTTGGGAAAGTACCA 1949
QY 1675 GATTGGTCAGAAACTGGGGGAGCCAAAGGTGCAACTCTCTATAATTTACTAGGTTCG 1734
Db 1950 GAATGGATGAAACTGGGGGAGCCTTAATATACAGAGGTATAAATTCACCGAGTTGC 2009
QY 1735 GCAGCTCACCATTCAGCACACGAAAGTACGCTCTCAGCCAGAACATGACACTAAT 1794
Db 2010 AAAGACT---TAGAGACACAAGCGCAAGCAATCCTCAGAGTCAAGACCAAGTTCTAAT 2066
QY 1795 CCACTTGCATCGATCTCAGGACCTGGCTTTAGAGCCTTGGAGCACACCAATATCTCT 1854
Db 2067 CCTCTGACTCCGGACGAGTGGGACCTTGCACCTGGAACCGTGGAGTACTCCAGATACGCT 2126
QY 1855 GTTGGCGGCACTGCAGAAACCCAGAACACCTGGGGAAGCTGGTCCAAAGCCTGCCAAGAT 1914
Db 2127 ATTGCAGAACTGCNAATCAACATCAACCAACTTGGCGTTACTCA-----CAAAGAC 2180
QY 1915 GGTCAACTGAGCCCAACTTGGTCAGAGATCGAGAGGATTTGAGAGCTGCTCGGTGCG 1974
Db 2181 GTGCAAGCGAGTCCGACGTGGTCCGAAATAGAGCAGACCTGAGAGCCATCTTTACTTCT 2240
QY 1975 GAACCGTTGAGAAAGACTTTTCAGGA 2000
Db 2241 GAACAATTTGGAAGAAGATTTTCGAGA 2266

RESULT 9
AAT88324 ID AAT88324 standard; DNA; 5049 BP.

AC AAT88324;

XX AC

DT 17-OCT-2003 (revised)

DT 21-MAY-1998 (first entry)

XX Attenuated canine parvovirus (vBI440) genomic DNA.

XX Canine parvovirus; CPV; attenuation; vBI440; vaccine; dog;
KW feline panleukopenia virus; mink enteritis virus; infection; ds.
XX Canine parvovirus; vBI440 (ATCC VR 2489) .

XX Key Location/Qualifiers
FH mutation 59
FT /tag= C
FT /note= "base 59 is G in CPV-39 (passage 5)"
FT mutation 97
FT /tag= d
FT /note= "base 97 is C or T in CPV-39 (passage 5)"
FT CDS 273..2279
FT /tag= a
FT /note= "NS1/NS2 coding region"
FT CDS 2286..4541
FT /tag= b
FT /note= "VP1/VP2 coding region"
FT mutation 4745
FT /tag= e
FT /note= "base 4745 is T in CPV-39 (passage 5)"
FT mutation 4881
FT /tag= f
FT /note= "base 4881 is C in CPV-39 (passage 5)"

XX W09742972-A1.

PN 20-NOV-1997.

XX 06-MAY-1997; 97WO-US007584.

XX 15-MAY-1996; 96US-00647655.

XX (CORR) CORNELL RES FOUND INC.

XX Parrish CR, Carmichael LE, Gruenberg A;

XX WPI; 1998-008583/01.

XX Canine parvovirus DNA carrying specific attenuating mutation(s) - used as
vaccines for protection against parvovirus and feline pan-leukopenia
virus infections.

XX Example 8; Page 34-37; 60pp; English.

XX This DNA sequence comprises an attenuated virus genome derived by serial
passaging (60 times) of virulent canine parvovirus (CPV) type 2b isolate
39 in NUPK feline kidney host cells. The attenuated virus is designated
vBI440 (ATCC VR 2489). It contains 4 mutations relative to the sequence
(see AAT88320) of the control (5th passage) wild-type CPV-2b. 2 Mutations
are within the hairpin formed by the 3' terminal palindromic: the mutation
at nucleotide 59 introduces an A into a G-C rich region within the tip of
the hairpin, disrupting the base pairing in one of the 2 small internal
palindromes within that sequence; the thymine at nucleotide 97 is
adjacent to the mismatched bubble (flip-flop) sequence within the
palindrome. The DNA from attenuated CPV strains (see also AAT88321) is
used for the production of infectious molecular DNA clones, which, in
turn, can be transfected into cells to generate master stocks of the
virus. The attenuated viruses can be used in dogs as a vaccine to protect
against CPV disease, or more generally in cats and minks to protect
against feline panleukopenia virus and mink enteritis virus. The vaccines
protect against the currently prevalent CPV-2b type (and all extant
strains of types 2 and 2a), providing a long term immune response.
(Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 5049 BP; 1791 A; 812 C; 1029 G; 1417 T; 0 U; 0 Other;

XX Query Match 51.5%; Score 1040.4; DB 2; Length 5049;

XX Best Local Similarity 71.4%; Pred. No. 8.9e-270;

XX Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;

QY 1 ATGCGTGGAAATGCTTACTCTGATGAGTTTGGGAGCAACCACTGTTAAAGGAAAAA 60

Db 273 ATGCTGGCAACCACTATCTAGGAAGTTATGAGGAGTAATGGTTAAGAAACAT 332
Qy 61 AGTAACAGGAGGTTCTCTCATTTGTTTTTAAAAATGAAATGTTCAACTGAATGGAAAA 120
Db 333 GCAGAAAATGAAGCATTTTCGTTGTTTTTAAATGTGACACGCTCCAACTAAATGGAAAG 392
Qy 121 GATATCGATGGAATAGTTACAAAAAGAGCTGAGGAGGAGCGAGCTGAATCTTTTACAA 180
Db 393 GATGTCGCTGGAAACAACATATACCAACCAATTCAAAATGAAGAGCTAAACATCTTTAAT 452
Qy 181 CGAGGAGCGGAACTACTTGGGACCAAGC-----GAGGACATGGAATGGAAACCA 234
Db 453 AGAGGAGCACHACAGCAATGGAATCAACCGAAGAGAGAAATGGACTGGGAATCGAA 512
Qy 235 GTGGATGAATGACCAAAAGCAAGTATTCATTTTGGATTTCTTTGGTTTAAAAATGTTTA 294
Db 513 GTTGATAGTCTCGCCAAAAGCAAGTACAACTTTTGTGATGATTAATTAATAATGTCCT 572
Qy 295 TTTGAAGTGTAAACACAGAAATATATTTCTCGGTGATGTTAATTTGGTTTGTGCAACAT 354
Db 573 TTTGAAGTCTTTGTTCTTAAAAATATAGAACCAATGAATGATGTTGTTTATTCACAT 632
Qy 355 GAATGGGGAAGACCAAGCTGCGCACTGCGCATCTACTTAATTTGAGGAAGAGCTTTAGT 414
Db 633 GAATGGGGAAGAGATCAAGGCTGGCATTTGTCATGTTTCTTTCATAGTAAGAACTTACAA 692
Qy 415 CAAGCTCAAGGGAATGGTGAGAAAGCAACTAAATGTTTACTGGGACAGATGTTGGTA 474
Db 693 CAAGCACTGGTAATGGCTAGCGAGACAAATGATATGATTTGGAGTAGATGGTTGGTG 752
Qy 475 ACAGCTGTAAATGCAACTAACACCAAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAA 534
Db 753 ACTCTTTGTTGCGTAAACTTAACACCAACTGAAAGATTAAGCTCAGAGAAATTCAGAA 812
Qy 535 GACATGAGTGGTTACTTACTTACTTATTAAGCATTAAGCAACCAAAAGAGCTATACC 594
Db 813 GATAGTGAATGGGTGACTATATTAACATACAGACATAAGCAACAACAAAAGAGACTATGTT 872
Qy 595 AAGTGTGTTCTTTTGGAAACATGATGCTTACTTATTTTAACTTAAAGAAATAAGC 654
Db 873 ABAATGTTCAATTTGGAAATATGATAGCATATTTACTTTTAAACAAAGAAATAATGTC 932
Qy 655 ACTAGTCCACAGAGACGGAGGTATTTTCTTAGCAGTGAATCTGCTGCGGAAACTAAC 714
Db 933 CACAT--GACAAAGAAAGTGGCTATTTTAAAGTACTGATTTCTGGTTGAAATTTAAC 989
Qy 715 TTTTAAAGAGCGGAGCGCATCTAGTAGCAAACTATACATGATGATGATGCGGCCA 774
Db 990 TTTATGAAGTATCAAGACAGACAAATTTGTGAGCACTTTTACCTGAACAAATGAAACCA 1049
Qy 775 GAAACGGTTGAAACCAACAGTAACCACTGCGAGGAAACTAAGCGCGGACAGAAATTCAAACT 834
Db 1050 GAAACGGTTGAAACCAACAGTACAGACAGACAGAAACAAAGCGCGGAGAAATTCAACT 1109
Qy 835 AAAAAGAAGTTTCTATTTAAACTACACTTAAAGAGCTGGTGCAATAAAGAGTAACTCA 894
Db 1110 AAAAAGGAAGTGTCAATCAATGTATTTTGGGAGCTTTGGTTAGTAAAGAGTAAACATCA 1169
Qy 895 CCAGAGACTGGATGATGATGACGCGACAGAGTTTACATTTGAATGATGGCTCAACAGGT 954
Db 1170 CCTGAAGACTGGATGATGTTACAAACCAATAGTATATTTGAATGATGGCAACACCAAGGA 1229
Qy 955 GGAGAAACCTGCTGAAAAATACCTAGAGATTTGTACACTAACTCTAGCCAGAACCAAA 1014
Db 1230 GGTGAAATCTTTTAAAAAATACACTTGAATTTGTACTTTGACTTTTAGCAAGAACAAA 1289
Qy 1015 ACAGCATTTGACTTAAATTTTAAAAAGAGCTGAAACCAAGCAACTTAACTTTTCACTG 1074
Db 1290 ACAGCATTTGAATTAATCTTTGAAAAAGCAGATATATACTAACTTAACTTATGATCTT 1349
Qy 1075 CCTGACACAAGACCTGACAGATTTTGTCTTTTATGCTGGAACATATGTTTAAAGTTTGC 1134

Db 1350 GCAAAATTTAGAACATGTCAAAATTTTGTAGAAATGACGGATGGAATTTGATTTAAAGTTTGT 1409
Qy 1135 CATGCTATTTGCTGTTGTTTAAACACAGACAAGGAGGCAAAAAGAAATACTGTTTATTTTCAT 1194
Db 1410 CACGCTATAGCATGTTGTTTTTAAATAGACAAGTGGTAAAGAAATATACAGTTCTTTTTTCAT 1469
Qy 1195 GGACCAAGCAGCAGCAAGCAATCTATTTATTTGCAAGCCATAGCACAAAGCAGTTGGCAAT 1254
Db 1470 GGACCAAGCAGTACAGAAATCTATCTATTTGCTCAAGCCATAGCACAAAGCTGTGGGTAA 1529
Qy 1255 GTTGGTTGCTATTAATGAGCAATGTAAACTTTTCCATTTTAAATGACTGTACCAACAAGAAC 1314
Db 1530 GTTGGTTGTTATTAATGAGCAATGTAAATTTTCCATTTTAAATGACTGTACCAATATAAAT 1589
Qy 1315 TTGATTTGGTAGAAGAGCTGGTAACTTTTGGACAGCAAGTAAACAGTTTAAAGCCATT 1374
Db 1590 TTAATTTGGATTGAAGAGCTGGTAACTTTTGGTCAACAACTTAATCAATTTTAAAGCAATC 1649
Qy 1375 TGCTCTGGTCAAGCTATTCGCAATTTGATCAAAAAGGAAAGGCAAGCAACAGATTTGAACCA 1434
Db 1650 TGTCTGGACAAACAAATTAGAATTTGATCAAAAAGGTAAGGAAGTAAGCAAAATTTGAACCA 1709
Qy 1435 ACACCACTCATGACCAACAAATGAGAACATTTACAGTGGTCAAGATAGGCTGCGAAGAA 1494
Db 1710 ACTCCAGTAATTAATGACAACTAATGAATAATAACAATTTGAGAAATTTGGATGTGAAGAA 1769
Qy 1495 AGACCAAGAACACTCAACCAATCAGACAGAGAAATGCTTAACTATCTAAACATACATACC 1554
Db 1770 AGACCTGAACATACACAAACCAATAAGACAGACAGAAATTTGAACATTTAAGTATGATGAAG 1829
Qy 1555 TTGCTGGTGACTTTGGTTGGTTGACAAAATGAATGGCCCATGATTTGCTGTTGCTG 1614
Db 1830 CTTCCAGGAGACTTTGGTTGGTTGATAAAGAAAGTGGCTTTTAAATATGTGCATGGTTA 1889
Qy 1615 GTAAAGATGGTTTACCAATCTACCATGGCAAGCTACTGTCTAAATGGGGCAAAAGTTTCCT 1674
Db 1890 GTTAACATGTTTGTGATCAACATGGCTTAACTATACATCATTTGGGGAAGATACCA 1949
Qy 1675 GATTGTCAGAAAACCTGGGCGGAGCCAAAGGTGCCAACTCCTATPAAATTTTACTAGGTTG 1734
Db 1950 GAATGGGATGAAAACCTGGGCGGAGCTTAAATAACAAGAGTATAAATTCACCAAGTTGC 2009
Qy 1735 GCAGCTCACATTCAGACACCGAAGAGTACGCTCTCAGCCAGAGACTATGCACACT 1794
Db 2010 AAAGACT---TAGAGACACAAGCGGCAAGCAATCTCAGAGTCAAGACCAAGTTCTAACT 2066
Qy 1795 CCATTTGATCGGATCTCGAGGACCTGGCTTTAGAGCTTTGGAGCAGACACCAAAATCTCCT 1854
Db 2067 CCTCTGACTCGGACGATGTGAGACCTTGCACTGGAAACCGTGGAGTACTCCAGATACGCT 2126
Qy 1855 GTTGGCGGCACTGCAAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAAGAT 1914
Db 2127 ATTGCAGAAACTGCAATCAACAAATCAACCAACTTGGCGTTACTCA-----CAAGAC 2180
Qy 1915 GGTCACTGAGCCCACTTTGGTCAGAGATCGAGAGAGATTTGAGAGCGTGTTCGGTGG 1974
Db 2181 GTGCAAGCGAGTCCGAGCTGGTCCGAAATAGAGGACAGCCTGAGAGCCATCTTTACTTCT 2240
Qy 1975 GAACCGTTTGAAGAAAGACTTTCAGCA 2000
Db 2241 GAACAAATTGGAAGAAGATTTTCGAGA 2266

RESULT 10

AAT88321
ID AAT88321 standard; DNA; 5049 BP.

XX AAT88321;

XX AC AC

XX 21-MAY-1998 (first entry)

XX Attenuated canine parvovirus genomic DNA.

XX

KW Canine parvovirus; CPV; attenuation; vaccine; dog;
KW feline panleukopenia virus; mink enteritis virus; infection; ds.

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	CDS	273.2279
----	-----	----------

FT /*tag= a

FT /note= "NS1/NS2 coding region"

FT	CDS	2286.	.454
----	-----	-------	------

FT /*tag= b

FT /note= "VP1/VP2 coding region"

FT mutation 4307

```
/*tag= c
```

/note= "base 4307 is A in virulent CPV-39 (G in passage

65 attenuated virus)¹¹

FT mutation 4358 / 4358

E

FT /note= "base 4358 is C in virulent CPV-39 (T in passage 35 constructed virus)"

FI	65 attenuated virus)
FT	4409
	mutation

FT
MUTATION
4409
/4433
0

FF /note=
FF /note=
FF /note=
FF /note=

FI Base 4403 IS C III Vincent CFV-33 (A III passage
/note= 65 attenuated vixing) "

FT mutation 4477

```

FT
macacorum
1177
/*tag= f

```

FT /note="ba

65 attenuated virus)"

FT mutation 4889

```
FT
/*tag=g
```

FT /note= "ba

FT 65 attenuated virus)"

FT mutation 4973

FT /*tag= h

FT /note= "ba

FT 65 attenuated virus)"

XX

PN WO9742972-A1.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

PD 20-NOV-1997.

XX
DE
06-MAY-1997. 07WA-115007594

06-MAY-1997; 9/WO-US00/384.
FF
YY

XX
PP 15-MAY-1996. 96UIS-00647655

FR 13-MAY-1990, 3003-XX

AA
PA
XX
(CORR) CORNELL RES FOUND INC.

PI Parrish CR, Carmichael LE,

XX
DR
XX
PT
PT
PT
PT
WFI; 1998-008583/01.
Canine parvovirus DNA carrying specific attenuating mutation(s) - used
vaccines for protection against parvovirus and feline pan-leukopenia
virus infections.

Db 1230 GGTGAAATCTTTTAAATAACACTTGAATTTGTACTTTGACTTTTAGCAAGACAAA 1289
Qy 1015 ACAGATTGCTTAATTTTAAAGAAAGCTGAACACCACTAACCAACTTTTCACTG 1074
Db 1290 ACAGATTGAAATTAATCTTGAAGAAAGCAGATAATACTAACTAACTTTTGATCTT 1349
Qy 1075 CTTGACACAGAACTGCAAGATTTTGTCTTTTATGCTGCTGGAATATGTTAAAGTTGC 1134
Db 1350 GCAAATCTGAAACATGTCAAATTTTGAAGATGACGAGTGAATTTGAAATTTTGT 1409
Qy 1135 CATCTATTTCTGTTTAAACAGACAGAGGCAAGAAATACCTGTTTAAATTTTCAAT 1194
Db 1410 CACCTATAGCAATGTTTAAATAGACAGAGTGTAAAGAAATACAGTCTTTTTCAT 1469
Qy 1195 GGACAGCAGACAGGCAATCTATTATTGCAAGCCATAGACAGCAAGTTGGCAAT 1254
Db 1470 GGACAGCAAGTACAGGAAATCTATCTTCTCAAGCCATAGCACAAGCTGTGGTAAT 1529
Qy 1255 GTTGGTTGCTATAGTCAAGCAATGTAACTTTTCCATTTAATGACTGTACCAAGAAC 1314
Db 1530 GTTGGTTGTTAATAGCAGCAATGTAAATTTTCCATTTAATGACTGTACCAATAAAT 1589
Qy 1315 TTGATTTGGGTAGAGAGCTGGTAACTTTTGACAGCAAGTAAACCACTTTAAAGCCATT 1374
Db 1590 TTAATTTGGATGAGAGCTGGTAACTTTTGTGCAAGATTAATCAATTTAAAGCAATC 1649
Qy 1375 TGCTCTGCTCAAGCTATTGCAATTTGATCAAAAGGAAAGGCAAGCAAGATTTGAACCA 1434
Db 1650 TGTTCTGCAAAACAAATTAGAAATTTGATCAAAAGGTAAGGAAGTAAAGCAATTTGAACCA 1709
Qy 1435 ACACAGTATCATGACCAACAAATGAGAAATTAACAGTGTGTCAGAAATAGGCTGGAAGAA 1494
Db 1710 ACTCAGTAATATGACAACTAATGAAATTTTAAATTTGTGAGAAATTTGATGTAAGAA 1769
Qy 1495 AGACCAAGACACACTCAACCAATCAGAGACAGAGTCTTAACATTTCACTAAACATACC 1554
Db 1770 AGACCTGAACATACAAACCAATAGAGACAGAAATTTGACATTAAGTATGATGTAAG 1829
Qy 1555 TTGCTGCTGACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1614
Db 1830 CTTCCAGGAGACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1889
Qy 1615 GTAAAGATGTTTACCAATCTACATGCGAAGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1674
Db 1890 GTTAAACATGTTTGAATCAACCATGCTGCTAACTATACATCATTTGGGAAAGTACCA 1949
Qy 1675 GATTGTCAGAAACTGGGCGGACCAAGTGCCCACTCTTAAATTTTACTAGTTGCG 1734
Db 1950 GAATGGATGAAACTGGGCGGACCTTAAATACAGAGGTATAAATTCACCAAGTTGC 2009
Qy 1735 GCAGCTCACCATTTACGACACCCGAAAGTACGCTCTCAGCCAGAACTATGCACATACT 1794
Db 2010 AAGACT---TAGAGACACAGCGGCAAGCAATCTCAGAGTCAAGACCAAGTTCTAAT 2066
Qy 1795 CCACTTCATCGGATCTGAGGACCTGGCTTTAGAGCCTTTGAGACACACCAATATCTCT 1854
Db 2067 CTTCTGACTCGGACGTTAGTGACCTTGCACTGCAACCGTGAGTACTCCAGATACGCT 2126
Qy 1855 GTTGGCGCACTGAGAAACCCAGAACACTGGGGAAGCTGGTTCCAAAGCCTGCCAGAT 1914
Db 2127 ATTGCAGAACTGCAAACTCAACCAATCAACCACTTGCGGTTACTCA-----CAAAGAC 2180
Qy 1915 GGTCAACTGAGCCCAACTTGGTCAGAGATGAGGAGGATTTGAGAGGCTGCTCGGTGCG 1974
Db 2181 GTGCAAGGAGTCCGAGCTGTTCCGAATAGAGGACACCTGAGGCCATCTTTACTTCT 2240
Qy 1975 GAACCGTTGAAGAAGACTTTACGGA 2000
Db 2241 GAACAAATTGGAAGAAGATTTTCGAGA 2266

RESULT 11
AAT88320

ID XX AAT88320 standard; DNA; 5049 BP.
XX AAT88320;
XX AC
XX 17-OCT-2003 (revised)
DT 21-MAY-1998 (first entry)
XX
XX Canine parvovirus 39 passage #5 (wild-type).
XX
XX Canine parvovirus; CPV; attenuation; vBI440; vaccine; dog;
KW feline panleukopenia virus; mink enteritis virus; infection; ds.
XX
XX Canine parvovirus; type 2b isolate 39.
XX
XX
XX Key Location/Qualifiers
CDS 273..2279
FT /*tag= a
FT /note= "NS1/NS2 coding region"
FT 2286..4541
FT /*tag= b
FT /note= "VP1/VP2 coding region"
XX
XX WO9742972-A1.
XX
XX 20-NOV-1997.
XX
XX 06-MAY-1997; 97WO-US007584.
XX
XX 15-MAY-1996; 96US-00647655.
XX
XX (CORR) CORNELL RES FOUND INC.
XX
XX Parrish CR, Carmichael LE, Gruenberg A;
XX
XX WPI; 1998-008583/01.
XX
XX Canine parvovirus DNA carrying specific attenuating mutation(s) - used as
PT vaccines for protection against parvovirus and feline pan-leukopenia
PT virus infections.
XX
XX Example 8; Page 37-40; 60pp; English.
XX
XX This DNA sequence comprises the genome of virulent canine parvovirus type
CC 2b (CPV-2b) isolate 39, obtained after 5 serial passages in NLFK feline
CC kidney host cells. Further passaging has yielded attenuated virus vBI440
CC (ATCC VR 2489) (see AAT88324) from the 60th passage, and a claimed virus
CC (see AAT88321) derived from the 65th passage (ATCC 2528). These
CC respectively contain 4 and 6 mutations in comparison to the virulent 5th
CC passage virus. The DNA from attenuated CPV-2b strains is used for the
CC production of infectious molecular DNA clones, which, in turn, can be
CC transduced into cells to generate master stocks of the virus. The
CC attenuated viruses can be used in dogs as a vaccine to protect against
CC CPV disease, or more generally in cats and minks to protect against
CC feline panleukopenia virus and mink enteritis virus. The vaccines protect
CC against the currently prevalent CPV-2b type (and all extant strains of
CC types 2 and 2a), and provide a long term immune response. (Updated on 17-
CC OCT-2003 to standardise OS field)
XX
SQ Sequence 5049 BP; 1789 A; 813 C; 1030 G; 1417 T; 0 U; 0 Other;
Query Match 51.5%; Score 1038.8; DB 2; Length 5049;
Best Local Similarity 71.3%; Pred. No. 2.4e-269;
Matches 1431; Conservative 0; Mismatches 557; Indels 18; Gaps 4;
Qy 1 ATGCTCGGAATGCTTACTCTGATGAGTTTTCGGAGCAACCACTGGTTAAAGGAAAA 60
Db 273 ATGCTCGCAACCAAGTATATCTGAGGAGTTATGCGGAGTAAATTTGTTAAAGAAACAT 332
Qy 61 AGTAACAGGAAGTGTCTCATTTGTTTAAATAATGAAATGTTCAACTGAATGAAAA 120
Db 333 GCAGAAATGAAGCATTTTCGTTGTTTAAATGTAAGCAACGTCCTCAACTTAATGGAAG 392
Qy 121 GATATCGGATGAATAGTTACAAAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTACAA 180

393	Db	GATGTTTCGCTGGAAACAACTATATACAAACCAATTCAAAAATGAAGAGCTAACATCTTTAAAT	452
181	Qy	CGAGGAGCGGAACACTACTCTGGGACCAAGC-----GAGGACATGGAATGGGAAACCA	234
453	Db	AGAGGAGCACAACAGCAATGGATCAACCGAAGAGAGAAATGGACTGGGAATCGGAA	512
235	Qy	GTGGATGAATGACCAAAAGCAAGTATTCATTTTTTGATCTTTGGTTAAAAAATGTTTA	294
513	Db	GTTGATAGTCTCGCCAAAGAGCAAGTACAACCTTTTGATGTCATTAATTAAGAAATGTC	572
295	Qy	TTTGAAGTCTTAACACAAAGATATATNTTCTGGTGAATGTTAATTTGGTTGTGCAACAT	354
573	Db	TTTGAAGTCTTTGTTCTAAAAATATAGAACCAATGATGTGTTGGTTTATATCAACAT	632
355	Qy	GAATGGGGAAAAGACCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAGGACTTTAGT	414
633	Db	GAATGGGGAAAAGATCAAGGCTGGCATTTGTCATGTTTACTTCATAGTAAGAACTTACAA	692
415	Qy	CAAGCTCAAGGGAAATGGTGGAGAAGCAACTAAATGTTTACTTTGGAGCAGATGGTTGGTA	474
693	Db	CAAGCAACTGGTAAATGGCTACCGAGACAATGAATATGTATTTGGAGTAGATGGTTGGTG	752
475	Qy	ACAGCCTGTAATGTGCAACTAACAACAGCTGAAAGAAATTAATTAAGAGAAATAGCAGAA	534
753	Db	ACTCTTTGTTGGTAAACTTTAAACCAACTGAAAGATTAAGCTCGAGAAATGCGAGAA	812
535	Qy	GACAACTAGCTGGGTACTCTACTTATTAAGCATTAAGCAACCAACCAAAAAAGGACTATACC	594
813	Db	GATAGTGAATGGGTGACTATATTAACATACAGACATAAGCAACCAAAAAAGGACTATGTT	872
595	Qy	AAGTGTCTCTTTTGGAAAACATGATGCTACTATTTTTTAACTAAAAAGAAAAATAAGC	654
873	Db	AAAATGGTTCATTTTGGAAATATGATAGCATATTAATTTTTTAACAAGAAAAAATTTGTC	932
655	Qy	ACTAGTCCAACCAAGAGACGGAGGCTATTTCTTAGCAGTCACTCTGGCTGGAAAACTAAC	714
933	Db	CACAT---GACAAAAAGAAAGTGGCTATTTTTTAAGTACTGATCTGGTTGCAAAATTAAC	989
715	Qy	TTTTTAAAGAGCGGAGCGCATCTAGTCAGCAAACTATACATGATGACATCGCGCCA	774
990	Db	TTTATGAAGTATCAAGACAGACAAATTTGTGAGCACACTTTACACTGAACAAATGAAACCA	1049
775	Qy	GAAACGGTGTGAACCAAGTAAACACATGCGCAGGAAACTAAGCGCGGACGAAATCAAACT	834
1050	Db	GAAACCGTTGAACCAAGTACAGCAGCAGGAAACAAAGCGCGGAGAAATCAAACCT	1109
835	Qy	AAAAAAGAAGTTCTATTAAGAACTACACTTTAAAGAGCTGTGTGCATATAAGAGTAACCTCA	894
1110	Db	AAAAAGGAAGTGTCAATCAATGTACTTTGCGGACTTGGTTAGTAAAGAGTAACATCA	1169
895	Qy	CCAGAGACTGGATGATGATGCAGCAGACAGTTTACATTTGAATGATGGCTCAACAGGT	954
1170	Db	CCTGAAGACTGGATGATGTTACAAACAGATAGTTTATATTGAATGATGGCAGCAACCGAGGA	1229
955	Qy	GGAGAAACCTGCTGAAAAATACGCTAGAGATTTGTACACTTAACCTAGCCAGAACCAAA	1014
1230	Db	GGTGAATAATCTTTTAAATAATACACTTTGAAATTTGTACTTTGACTTTAGCAAGAACAAA	1289
1015	Qy	ACAGCATTTGACTTAATTTTAGAAAAAGCTGAAACAGCAAACTTAACCACTTTTCACTG	1074
1290	Db	ACAGCATTTGAATTAATCTTTGAAAAAGCAGATTAATACTAAACTTAACCTTTGATCTTT	1349
1075	Qy	CCTGACACAGAACCTGCAGAAATTTTTTTCATCTGGCTGGAACTATGTTAAGTTTGC	1134
1350	Db	GCAAAATCTAGAACATGTCAAAATTTTTTAGAATGCAACGGATGGAATTTGGAATTAAGTTTGT	1409
1135	Qy	CATGCTATTTGCTGTGTTTTTAAACAGACAAAGGAGCAAAAGAAATACGTGTTTATTTTCAT	1194
1410	Db	CACGCTATGACATGTGTTTTTAATATAGACAGGTGGTAAAGAAATACAGTTCCTTTTCAT	1469
1195	Qy	GGACCAGCCAGCACAGGCAATCTATTTATGCAAGCCATAGCACAAAGCAGTTGGCAAT	1254

Db	1470	GGACCAAGATGACGAGAAATCTATCATTGCTCAGCCATAGCACAAAGCTGTGGGTAA	1521
Qy	1255	GTGTGGTTGGCTATAAATGCAGCCAAATGTAAACTTTTCCATTTAAATGACGTATCCAAACAAGAAC	1314
Db	1530	GTGTGGTTGTATTAATGCAGCAAAATGTAAATTTTCCATTTAAATGACGTATCCAAATAAAAAT	1589
Qy	1315	TTGATTTGGGTGAGAAGAGCTGTGTAACTTTTGGGACGCAAGTAAACCAAGTTTAAAGCCATT	1374
Db	1590	TTAAATTTGGATTGAAGAAGCTGTGTAACTTTTGGTCAACAAGTTTAACTCAATTTTAAAGCAATC	1649
Qy	1375	TGCTCTGGTCAAGTATTGCGATTGTATCAAAAAGGAAAAAGGACGAGCAAAACAGATTGTAACCA	1434
Db	1650	TGTTCTTGGACAAACAATTTAGATTGTATCAAAAAGGTTAAAGGAAAGTAAAGCAAAATGTAACCA	1709
Qy	1435	ACACCAAGTCAATCATAGCCACAAATGAGAACATTTACAGTGGTCAAGATAGAGCTCGCAAGAA	1494
Db	1710	ACTCCAGTAATTTATGACAACTAATGAAATTTTAAACAAATTTGTGAGAATTTGGATGTGAAGAA	1769
Qy	1495	AGACCAAGACACACTCAACCAATTCAGAGACAGAAATGCTTTAAACATTTATCTAAACACATATACC	1554
Db	1770	AGACCTGAACATACACAAACCAATAAGAGAGACAGAAATGTTGAACATTTAAGTTAGTATGTGAAG	1829
Qy	1555	TTGCCTGGTGACATTTGGTTTGGTTGTACAAAAATGAATGGCCCAATGATTTCTGCTTGGTTTG	1614
Db	1830	CTTCCAGGAGACATTTGGTTTGGTTGTATAAGAAAGAAATGGCCCTTTAATATGTGCATGGTTA	1889
Qy	1615	GTAAAGAATGGTTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGCGCAAAAGTTCCCT	1674
Db	1890	GTTTAAATGTTTGTGAAATCAACCATGGCTTAACTATACACATCATTTGGGGAAGTACCA	1949
Qy	1675	GATTGTGAAAAACTTGGGCGGAGCCAAAGGTGCCAACTCTCTATAAATTTACTAGGTTGCG	1734
Db	1950	GAATGGGATGAAAACCTGGGCGGAGCCCTTAAATAACAAGAAGGTATAAATTCACCAGGTTGC	2009
Qy	1735	GCAGCTCACCATTACGACACCCGAAAGGTACGCCCTCTCAGCCAGAACTATGCACTAACT	1794
Db	2010	AAAGACT--TAGAGACAACAGCGGCAAGCAATCTCTCAGAGTCAAGACCAAGTTCTAACT	2066
Qy	1795	CCACTTGCATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGAGAGCCTTGGAGCACACCAAAATCTCCT	1854
Db	2067	CCTCTGACTCCGACGCTAGTGACCTTGCACTGGAACCGTGGAGTACTCCAGATACGCCT	2126
Qy	1855	GTTCGGGCACTGCGAGAAACCCGAGACACATGCGGGAGAGCTGGTTCCAAAGCCTGCCANGAT	1914
Db	2127	ATTGCAAAAATCGAAATCAACAATCAAACTTTGGCGTTACTCA-----CAAAGAC	2180
Qy	1915	GGTCAACTGAGCCCACTTGGTCAGAGATCGAGGAGATTGTAGAGCGTCTTCGGTGGCG	1974
Db	2181	GTGCAACGAGTCCGACGTGGTCCGAAATAGAGGCGAGCCCTGAGAGCCATCTTTACTTCT	2240
Qy	1975	GAACCGTTGAAGAAAGACTTTCAGCGA	2000
Db	2241	GAACAATTTGGAAGAGATTTCGAGA	2266
RESULT 12			
AAN40252			
ID	AAN40252 standard; DNA; 3524 BP.		
XX	AAN40252;		
AC			
XX			
XX			
DT	(revised)		
DT	24-OCT-2003		
DT	12-JAN-1992 (first entry)		
XX			
XX			
DE	Sequence from the double-stranded replicative form DNA of porcine parvovirus.		
DE			
XX			
KW	Protein envelope; immunogen; vaccine; antigen; epitope; ds.		
XX			
OS	Porcine parvovirus; NADL-2 virulent strain.		
XX			
Key	Location/Qualifiers		
FH	1..2073		
FT	CDS		

1567	CCAGGTGATTTTGACATTTT	TAGAGAACTGAATGGCCCACTAATATATGTGCTTGGTTGGTA	1628
1618	AAGAATGGTTACCAATCTACCATCGCAAGCTACTGTGCTAAATGGGCAAAAGTTCTCTGAT	1677	
1627	AAGAAAGGTTACCAAGCAACAATGGCTAGCTATATGTCATCATTTGGGGAATATGTAACCTGAT	1686	
1678	TGGTCAGAAAACCTGGGCGGAGCCAAAGTGCCAACTCCTATTAATTTTACTTAGGTTTCGGCA	1737	
1687	TGGTCAGAAAAAATTTGGAGGAGCCAAAATATGCATTTCCCAATAAATAACCA---ACAGAC	1743	
1738	CGCTCACCATTTACGACACCGGAAAAGTAGTCGGCTCTCAGCCAGAACTATGCACTAACTCCA	1797	
1744	TCTCAGATTTCCACATCAGTGAAAACCTTTGGCCGGGACATCACTACGACGCACTCCA	1803	
1798	CTTGCAATCGGATCTCGA---GGA	CCTGGCTTTAGAGCCTTTGGAGCAACCAAAATATCTCC	1853
1804	ATACAGAGGACCTTGGATTTAGCTTTTAGCTTTAGCGCTTTGGAGCGCTTTGGAGCGGAGCCAAACACAC	1862	

RESULT 13	
ABQ95626	
ID	ABQ95626 standard; DNA; 374 BP.
XX	
XX	
AC	ABQ95626;
XX	
XX	
DT	28-OCT-2002 (first entry)
XX	
XX	
DE	Tumour suppression-related oligonucleotide #1277.
XX	
KW	Tumour; cytostatic; antiviral; neuroprotective; nootropic; neuroleptic;
KW	tumour suppression; tumour reversion; apoptosis; viral resistance; human;
KW	viral infection; cell degeneration disease; neurodegeneration; ds;
KW	Alzheimer's disease; schizophrenia; immune disease; inflammatory disease.

The present invention relates to novel human nucleic acid sequences (I). The present sequence is one such nucleic acid sequence. Expression of (I) are implicated in tumour suppression or reversion and apoptosis and viral resistance. (I) are useful as probes or primers for detecting, identifying, measuring and/or amplifying nucleic acid sequences, as antisense reagents and for recombinant production of polypeptides. (I), polypeptides (II) encoded by (I), vector containing (I), cells containing these vectors and antibodies (Ab) against (II) are all useful for treatment/prevention of viral, tumour and cell degeneration diseases (especially neurodegeneration, such as Alzheimer's disease and schizophrenia). Analysing the expression of (I) is also useful for diagnosis and/or prognosis of such diseases. Transgenic animals carrying (I) are used for studying the aetiology of these diseases (also immune and inflammatory diseases). Note: in the present specification, SEQ ID 1 to 2280 are claimed in Claim 1, however only SEQ ID 1 to 2270 are shown in the specification

5Q	Sequence	374 BP; 116 A; 57 C; 108 G; 84 T; 0 U; 9 Other;	
	Query Match	15.2%; Score 307.4; DB 6; Length 374;	
	Best Local Similarity	88.2%; Pred. No. 2e-72;	
	Matches 329; Conservative	0; Mismatches 44; Indels 0; Gaps 0;	
Qy	174	TTTACAACGAGGAGCGGAACTACTTGGGACCAAGCGAGGACATGGGAATGGAAACCCAC	233
Db	1	TTTACAACGAGGCGGARGACCACTTGGGACCAAGCGAGGACATGGGAATGGGAGAGCGC	60
Qy	234	AGTGGATGAAATGACCAAAAAAGCAAGTATTTCATTTTTCGATCTTTTGGTTAAAAAATGTTT	293
Db	61	AGTGGATGACATGACCAAAAAAGCAAGTATTTCATTTTTCGATCTTTTGGTTAAAGAAGTGT	120
Qy	294	ATTTGAAGTGTCTTAACCAAGAAATATATTTCTGTGGTGATGTTAATTCGTTTGTGCAACA	353
Db	121	GTTTGAAGTGTCTAGCAACAAGAACAATAGCTCTTAGTAAATGTTACTTGGTTCGTGCAGCA	180
Qy	354	TGAATGGGAAAAAGACCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAGGACTTTTAG	413
Db	181	TGAATGGGAAAAAGGACCAAGGCTGGCACTGTTCATGTCTGATTTGGAGGCAAGGACTTTAG	240
Qy	414	TCRAGCTCAAGGGAAATGGTGGGAAGGCGCACTAAATCTTTTACTTGGACAGATGTTGCT	473
Db	241	TCRACCTCAAGGAAAATGGTGGGAAGGCGAGCTAAATGTGTACTGGAATANAATGTTTGN	300
Qy	474	AACAGCCTGTAATGTGCAACTTAACACCAAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGA	533
Db	301	GACTGCTCTGNAATGTGCAACTTAACACCAAGCTGAAAGAAATTAACCTGNNAGAAATAGCAGA	360
Qy	534	AGCAATAGCTGG 546	
Db	361	GGACAGTGANNNG 373	
RESULT 14			
ABQ94779			
ID	ABQ94779 standard; DNA; 421 BP.		
XX	ABQ94779;		
XX	28-OCT-2002 (first entry)		
XX	Tumour suppression-related oligonucleotide #430.		

RESULT 14	
ABQ94779	
ID	ABQ94779 standard; DNA; 421 BP.
XX	
XX	
AC	ABQ94779;
XX	
XX	
DT	28-OCT-2002 (first entry)
XX	
DE	Tumour suppression-related oligonucleotide #430.
XX	
XX	
KW	Tumour; cytostatic; antiviral; neuroprotective; nootropic; neuroleptic;
KW	tumour suppression; tumour reversion; apoptosis; viral resistance; human;
KW	viral infection; cell degeneration disease; neurodegeneration; ds;
KW	Alzheimer's disease; schizophrenia; immune disease; inflammatory disease.
XX	
XX	
OS	Homo sapiens.
XX	
XX	
PN	FR2819824-A1.
XX	
PD	26-JUL-2002.
XX	
XX	
PF	23-JAN-2001; 2001FR-00000899.
XX	
XX	
PR	23-JAN-2001; 2001FR-00000899.
XX	
XX	
PA	(MOLE-) MOLECULAR ENGINES LAB SA.
XX	
PI	Teclerman A, Amson R, Tuijnder M, Susini L;
XX	
XX	WPI; 2002-610803/66.
XX	
XX	
PT	New nucleic acid implicated e.g. in tumor suppression, useful for
PT	diagnosis of tumors, viral infection and cellular degeneration and for
PT	drug screening.
XX	
XX	
PS	Claim 1; Page 141; 623pp; French.
XX	
XX	
CC	The present invention relates to novel human nucleic acid sequences (I).

CC The present sequence is one such nucleic acid sequence. Expression of (I)
CC are implicated in tumour suppression or reversion and apoptosis and viral
CC resistance. (I) are useful as probes or primers for detecting,
CC identifying, measuring and/or amplifying nucleic acid sequences, as
CC antisense reagents and for recombinant production of polypeptides. (I),
CC polypeptides (II) encoded by (I), vector containing (I), cells containing
CC these vectors and antibodies (Ab) against (II) are all useful for
CC treatment/prevention of viral, tumour and cell degeneration diseases
CC (especially neurodegeneration, such as Alzheimer's disease and
CC schizophrenia). Analysing the expression of (I) is also useful for
CC diagnosis and/or prognosis of such diseases. Transgenic animals carrying
CC (I) are used for studying the aetiology of these diseases (also immune
CC and inflammatory diseases). Note: In the present specification, SEQ ID 1
CC to 2280 are claimed in Claim 1, however only SEQ ID 1 to 2270 are shown
XX in the specification
SQ Sequence 421 BP; 125 A; 77 C; 118 G; 96 T; 0 U; 5 Other;

Query Match 14.2%; Score 285.8; DB 6; Length 421;
Best Local Similarity 87.6%; Pred. No. 1.4e-66;
Matches 311; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 173 CTTTACACGAGGCGGAACTACTTGGACCAAGCGAGGACATGGAATGGAAACCA 232
DB 59 CTTTACACGAGGCGGAGACCACTTGGACCAAGCGAGGACATGGAATGGAGCG 118
QY 233 CAGTGGATGAATGACCAAAAGCAAGTATTTCATTTTGGTTTAAAAAATGTT 292
DB 119 CAGTGGATGACATGACCAAAAGCAAGTATTTCATTTTGGTTTAAAAAATGTT 178
QY 293 TATTTGAAGTGCTTAACACAAAGATATATTTCTGTGTATGTTAATTTGTTGCAAC 352
DB 179 TGTTTGAAGTGCTCAGCACAAAGAACATAGCTCTAGTATGTTTACTTGTGTCGAGC 238
QY 353 ATGAATGGGAAAGACCAAGGCTGGCAGTCTGCTACTTAATTTGGAGGAAGGACTTTA 412
DB 239 ATGAATGGGAAAGACCAAGGCTGGCAGTCTGCTACTTAATTTGGAGGAAGGACTTTA 298
QY 413 GTCAAGCTCAAGGAAATGGTGAGAGGCAACTAAATGTTTACTTGGAGCAGATGTTGG 472
DB 299 GTCAACCTCAAGGAAATGGTGAGAGGCAACTAAATGTTTACTTGGAGTAGATGTTGG 358
QY 473 TAACAGCTGTATGTGCAACTAACACGCTGAAGAAATTAACCTAAGAGAAAT 527
DB 359 TGACTGCTGTATGTTCACTAACACGCTGAAGAAATTAACCTAAGAGAAAT 413

RESULT 15

AB094724
ID AB094724 standard; DNA; 423 BP.

AC AB094724;

XX 28-OCT-2002 (first entry)

DE Tumour suppression-related oligonucleotide #375.

XX Tumour; cytostatic; antiviral; neuroprotective; nootropic; neuroleptic;
KW tumour suppression; tumour reversion; apoptosis; viral resistance; human;
KW viral infection; cell degeneration disease; neurodegeneration; ds;
KW Alzheimer's disease; schizophrenia; immune disease; inflammatory disease.

OS Homo sapiens.

XX FR2819824-A1.

XX 26-JUL-2002.

XX 23-JAN-2001; 2001FR-00000899.

XX 23-JAN-2001; 2001FR-00000899.

XX (MOLE-) MOLECULAR ENGINES LAB SA.

XX

PI Telerman A, Anson R, Tuijnder M, Susini L;

XX WPI; 2002-610803/66.

XX New nucleic acid implicated e.g. in tumor suppression, useful for
PT diagnosis of tumors, viral infection and cellular degeneration and for
PT drug screening.

XX Claim 1; Page 125-126; 623pp; French.

XX The present invention relates to novel human nucleic acid sequences (I).
XX The present sequence is one such nucleic acid sequence. Expression of (I)
CC are implicated in tumour suppression or reversion and apoptosis and viral
CC resistance. (I) are useful as probes or primers for detecting,
CC identifying, measuring and/or amplifying nucleic acid sequences, as
CC antisense reagents and for recombinant production of polypeptides. (I),
CC polypeptides (II) encoded by (I), vector containing (I), cells containing
CC these vectors and antibodies (Ab) against (II) are all useful for
CC treatment/prevention of viral, tumour and cell degeneration diseases
CC (especially neurodegeneration, such as Alzheimer's disease and
CC schizophrenia). Analysing the expression of (I) is also useful for
CC diagnosis and/or prognosis of such diseases. Transgenic animals carrying
CC (I) are used for studying the aetiology of these diseases (also immune
CC and inflammatory diseases). Note: In the present specification, SEQ ID 1
CC to 2280 are claimed in Claim 1, however only SEQ ID 1 to 2270 are shown
XX in the specification
SQ Sequence 423 BP; 125 A; 74 C; 124 G; 99 T; 0 U; 1 Other;

Query Match 14.1%; Score 284.8; DB 6; Length 423;

Best Local Similarity 89.2%; Pred. No. 2.7e-66;

Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 173 CTTTACACGAGGCGGAACTACTTGGACCAAGCGAGGACATGGAATGGAAACCA 232

DB 71 CTTTACACGAGGCGGAGACCACTTGGACCAAGCGAGGACATGGAATGGAGAGCG 130

QY 233 CAGTGGATGAATGACCAAAAGCAAGTATTTCATTTTGGTTTAAAAAATGTT 292

DB 131 CAGTGGATGACATGACCAAAAGCAAGTATTTCATTTTGGTTTAAAAAATGTT 190

QY 293 TATTTGAAGTGCTTAACACAAAGATATATTTCTGTGTATGTTAATTTGTTGCAAC 352

DB 191 TGTTTGAAGTGCTCAGCACAAAGACATAGCTCTAGTATGTTTACTTGTGTCAGC 250

QY 353 ATGAATGGGAAAGACCAAGGCTGGCAGTCTGCTACTTAATTTGGAGGAAGGACTTTA 412

DB 251 ATGAATGGGAAAGACCAAGGCTGGCAGTCTGCTACTTAATTTGGAGGAAGGACTTTA 310

QY 413 GTCAAGCTCAAGGAAATGGTGAGAGGCAACTAAATGTTTACTTGGAGCAGATGTTGG 472

DB 311 GTCAACCTCAAGGAAATGGTGAGAGGCAACTAAATGTTTACTTGGAGTAGATGTTGG 370

QY 473 TAACAGCTGTATGTGCAACTAACCCAGCTGAAGAAATTA 516

DB 371 TGACTGCTGTATGTTCACTAACCCAGCTGAAGAAATTA 414

Search completed: January 22, 2005, 11:50:01

Job time : 957.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 19, 2005, 16:54:30 ; Search time 164.5 Seconds
(without alignments)
8723.907 Million cell updates/sec

Title: US-10-069-056-16

Perfect score: 2019

Sequence: 1 atggctgggaatgcttactc.....agccgctgaactggactaa 2019

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 35539441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*

5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1040.4	51.5	5049	1	US-08-336-345-1
2	1040.4	51.5	5049	1	US-08-336-345-2
3	1040.4	51.5	5049	2	US-08-647-655-1
4	1040.4	51.5	5049	2	US-08-647-655-2
5	107.8	5.3	4680	1	US-08-254-358-1
6	107.8	5.3	4680	1	US-08-475-391-1
7	107.8	5.3	4680	2	US-08-709-609-1
8	107.8	5.3	4680	5	PCT-US95-07178-1
9	107.8	5.3	4681	4	US-08-807-802A-18
10	107.8	5.3	4910	2	US-08-331-384-2
11	107.8	5.3	4910	2	US-08-836-087-2
12	107.8	5.3	4910	3	US-09-246-320-2
13	107.8	5.3	4910	3	US-09-546-738-2
14	107.8	5.3	7214	4	US-09-438-268-1
15	107.8	5.3	7557	4	US-09-770-315-3
16	107.8	5.3	8151	4	US-09-438-268-2
17	107.8	5.3	8179	4	US-09-438-268-5
18	107.8	5.3	8698	4	US-09-770-315-2
19	103	5.1	939	4	US-09-532-594B-12
20	103	5.1	1197	4	US-09-532-594B-13
21	103	5.1	1611	4	US-09-532-594B-14
22	103	5.1	1872	4	US-09-532-594B-3
23	103	5.1	1872	4	US-09-532-594B-15
24	103	5.1	4767	4	US-09-532-594B-1
25	98.2	4.9	969	4	US-08-807-802A-10
26	98.2	4.9	1200	4	US-08-807-802A-8
27	98.2	4.9	1641	4	US-08-807-802A-6

Qy

1 ATGCGTGGAAATGCTTACTCTGATGAGTTTGGGACCAACCACTGTTAAGGAAAAA 60
Query Match 51.5%; Score 1040.4; DB 1; Length 5049;
Best Local Similarity 71.4%; Pred. No. 1.2e-292;
Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;

ALIGNMENTS

RESULT 1

US-08-336-345-1

; Sequence 1, Application US/08336345

; Patent No. 5814510

; GENERAL INFORMATION:

; APPLICANT: Parrish, Colin R.

; APPLICANT: Gruenberg, Allen

; APPLICANT: Carmichael, Leland E.

; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: NY

; ZIP: 10036-2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/336,345

; FILING DATE:

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Gordon, Jennifer

; REGISTRATION NUMBER: 30753

; REFERENCE/DOCKET NUMBER: 7937-006

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 790-9090

; TELEFAX: (212) 869-8864/9741

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5049 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: both

; TOPOLOGY: unknown

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; ORGANISM: Parvovirus

; US-08-336-345-1

Sequence 4, Appli
Sequence 19, Appli
Sequence 1, Appli
Sequence 14, Appli
Sequence 22, Appli
Sequence 22, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 2813, Ap
Sequence 14265, A
Sequence 1, Appli
Sequence 1, Appli
Sequence 325, App
Sequence 8976, Ap

Db 273 ATGCTGGCAACCCAGTATATCTGAGGAAGTTATGGAGGGAGTAAATGGTTAAAGAAACAT 332
Qy 61 AGTAAACGAGGAAGTGTCTCATTTGTTTAAATAAGAAATGTTCAAGCTGAATGGAARA 120
Db 333 GCAGAAAAATGAACATTTTGGTTGTTTAAATGTGACAACTGCTCAATTAATGGAAG 392
Qy 121 GATATCGGATGGAATAGTTACAAAAAGAGCTCGAGGAGCAGCTGAAATCTTTTACAA 180
Db 393 GATGTTGCTGGAAACAACTATACAAACCAATTCAAAATGAAGAGCTAACATCTTTAAT 452
Qy 181 CGAGGAGCGGAATCTACTTTGGGACCAAGC-----GAGGACATGAAATGGGAAACCA 234
Db 453 AGAGGAGCACAACAGCAATGGATCAACCGAAGAGAGAAATGACTGGGAATCGGAA 512
Qy 235 GTGGATGAATGACCAAAAAAGCAAGTATTCATTTTGGTTGTTGTTTAAATAATGTTTA 294
Db 513 GTTGATAGTCTCGCCAAAAAGCAAGTACAAACTTTTGTGATCAATTAATTAATAATGTC 572
Qy 295 TTTGAAGTGTCTTAACACAAAGAAATATATTTCTGGTGATGTTAATGGTTTGTGCAACAT 354
Db 573 TTTGAAGTCTTTGTTCTAAAAATATAGAACCAATGATGTGTTGGTTTATTCAACAT 632
Qy 355 GAAATGGGAAAGAACCAAGCTGGCACTGCGATGTACTAATTTGGAGGAAGGACTTTAGT 414
Db 633 GAATGGGAAAGATCAAGCTGGCAATGTCAATGTTTACTTCTATAGTAAGAACTTACAA 692
Qy 415 CAAGCTCAAGGGAATGGTGGAGAGCAACTAAATGTTTACTTGGAGCAGATGGTTGGTA 474
Db 693 CAAGCAACTGGTAAATGGCTACGAGACAAATGAATGATTTGGAGTAGATGGTTGGTG 752
Qy 475 ACAGCTCTAATGTGCAACTAAACACCACTGGAAGAAATTAACCTAAGAGAAATAGCAGAA 534
Db 753 ACTCTTTGTTGCGTAAACTTTAACACCAACTGAAAGATTAAGCTCAGAGAAATGCGAA 812
Qy 535 GACAATGAGTGGGTTACTCTACTTATTAAGCATTAAGCAACCAACCAAGAAAGACTATACC 594
Db 813 GATAGTGAATGGGTGACTATATAACATACAGACATAAGCAAAACAAAAAGAACTATGTT 872
Qy 595 AAGTGTCTCTTTTGGAAACATGATTGCTTACTTATTTTAACTAAAAAGAAATTAAGC 654
Db 873 AAAATGGTTCTATTTTGGAAATATGATAGCATATTACTTTTAAAGAAATAATGTC 932
Qy 655 ACTAGTCCACCAAGACGCGAGGCTATTTTCTTAGCAGTGACTCTGCTGGAAACCTAAC 714
Db 933 CACAT---GACAAAGAAAGTGGCTATTTTAAAGTACTGATTCGTTGGAAATTTAAC 989
Qy 715 TTTTAAAGAGCGGAGCGCCATCTAGTGAGCAACTATACACTGATGATGACATGCGGCCA 774
Db 990 TTTATGAAGTATCAAGACAGACAAATTTGTCAAGCACCTTTTACACTGAACAAATGAACCA 1049
Qy 775 GAAACGGTTGAAACCAACAGTAACCACTCGCAGGAAACTAAGCGCGCAGAAATTCAAACT 834
Db 1050 GAAACCGTTGAAACCAACAGTGAACAGACAGCAAGAAACAAAGCGCGGAGNAATCAACT 1109
Qy 835 AAAAAGAAAGTTTCTTATTAACACTACACTTAAAGAGCTGGTGCAATAAAGAGTAACCTCA 894
Db 1110 AAAAAGGAAGTGTCAATCAATGTACTTTTGGGACCTTGGTTAGTAAAGAGTAACATCA 1169
Qy 895 CCAGAGGACTGGATGATGATCGCCGACAGTGTACATTTGAATGATGGCTCAACACAGGT 954
Db 1170 CCTGAAGACTGGATGATGTTACAAACCAAGATAGTTATTTGAATGATGGCAACACCAAGGA 1229
Qy 955 GGAGAAAACTGCTGAAAAATACGCTAGAGATTTTGTACACTAATCTAGCCAGAACCAAA 1014
Db 1230 GGTGAATCTTTTAAAAAATACACTTGAAATTTGTACTTTTGAATTTAGCAAGAACCAAA 1289
Qy 1015 ACAGATTGACTTAAATTTTGAAGAAAGCTGAACACCAAGCAACTAACCACTTTCACTG 1074
Db 1290 ACAGCATTTGAATTAATCTTTGAAAAAGCAGATAATACTAAACTAACTAACTTTGATCTT 1349
Qy 1075 CTTGACACAAGAACTCGAGAAATTTTCTTTTCTGCTGCAACTATGTTTAAAGTTTGC 1134

Db 1350 GCAAAATCTAGAAATGTCAAATTTTTAGAAATGCACGGATGGAATTTGGATTAAAGTTGT 1409
Qy 1135 CATGCTATTTGCTGTGTTTTTAAACAGACAGAGGCAAAAAGAAATATCTGTTTTATTTCAT 1194
Db 1410 CACGCTATAGCATGTGTTTTTAAATAGACAGGTGGTAAAGAAATACAGTTCTTTTTCAT 1469
Qy 1195 GACACGACGACAGCAGGCAAAATCTATTATTGACAAAGCCATAGCACAGCAGTTCGCAAT 1254
Db 1470 GGACCAAGCAAGTACAGGAAATCTATCATTTGCTCAAGCCATAGCACAAAGCTGTGGGTAT 1529
Qy 1255 GTTGTGTGCTATAATGTCAGCCAAATGTAACCTTTTCATTTAATGACTGTACCAACAAAGAAC 1314
Db 1530 GTTGTGTGTTATTAATGTCAGCAATGTAATTTTCCATTTAATGACTGTACCAATAAATA 1589
Qy 1315 TTGATTTGGGTAGAAAGAGCTGGTAACTTTTGACAGCAAGTAAACAGTTTAAAGCCAT 1374
Db 1590 TTAATTTGGATTGAAGAGCTGGTAACTTTTGGTCAACAAAGTTAACTCAATTTAAAGCAATC 1649
Qy 1375 TGCTCTGTCAGCTATTTCGCAATTTGATCAAAAGGAAAGGAGCAGCAACAGATTGAACCA 1434
Db 1650 TGTCTCGGACAAACCAATTAGAATTTGATCAAAAGGTAAGGAAGTAAGCAAAATTTGAACCA 1709
Qy 1435 ACACCAAGTCAATCATGACCAACCAATGAGAACATTAACAGTGGTCAGAAATAGGCTGCGAAGAA 1494
Db 1710 ACTCCAGTAAATTTATGCAACTTAATGAAATATAACAAATTTGTGAGAATTTGATGTGAAGAA 1769
Qy 1495 AGACCAAGAACACACTCAACCAATCAGACAGACAGAAATGCTTTAACTTCACTTAACATACC 1554
Db 1770 AGACCTGAACATACACCAACCAATTAAGAGACAGAAATGTTGAACATTTAAGTTAGTATGTAAG 1829
Qy 1555 TTGCTCTGCTGACTTTGTTGTTGTCACAAAATGAATGGCCCATGATTTGTGCTGGTTG 1614
Db 1830 CTTCCAGGAGACTTTGTTGTTGTTGATTAAGAAAGTAAGCCCTTAAATATGTCATGTTA 1889
Qy 1615 GTAAAGAAATGTTTACCAATCTACCATGGCAAGCTACTGTGCTAAATGGGGCAAGTTCTCT 1674
Db 1890 GTTAAACATGTTTGAATCAACCAATGGCTAACTATACACATCATTTGGGGAAAAAGTACCA 1949
Qy 1675 GATTGTGCAAGAACTGGGGGAGCAGCAAGGTGCCAACTCTCTATAAATTTACTAGTTCTG 1734
Db 1950 GAATGGATGAATACTGGGCGAGCCCTAAATATACAAAGAGGTATAAATTCACAGGTTGC 2009
Qy 1735 GCAGCTCACCAATTCACGACACCGAAAAAGTAGCCCTCTCAGCCAGAACTATGCACTAAC 1794
Db 2010 AAAGACT---TAGAGACAAAGCGGCAAGCAATCTCTCAGAGTCAAGCCAAAGTTCTAACT 2066
Qy 1795 CCATTTGCTATCGGATCTCGAGGACCTGGCTTTTAGAGCTTTGAGGACACACAAATATCTCCT 1854
Db 2067 CCTCTGACTCCGACCGTAGTGGACCTTGCACTGGAAACCGTGGAGTACTCCAGATACGCT 2126
Qy 1855 GTTGGCGGCACTGCAGAAACCCAGAACACTGGGGAAGCTGTTCCAAAAGCTGCCCAAGAT 1914
Db 2127 ATTGCAGAACTGCAAAATCAACAACTCAACAACTTTGGCGTTACTCA-----CAAAGAC 2180
Qy 1915 GGTCAACTGAGCCCAACTTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTGTCTTCGGTGG 1974
Db 2181 GTCGAACGAGTCCGAGCTGGTCCGAATAGAGGCAAGCTGAGAGCCATCTTTTACTTCT 2240
Qy 1975 GAACCGTTGAAGAAAGACTTTCAGCA 2000
Db 2241 GAACAAATTTGAAGAAAGATTTTCGAGA 2266

RESULT 2
US-08-336-345-2
; Sequence 2, Application US/08336345
; Patent No. 5814510
; GENERAL INFORMATION:
; APPLICANT: Parrish, Colin R.
; APPLICANT: Gruenberg, Allen
; APPLICANT: Carmichael, Leland E.
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,345
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Gordon, Jennifer
; REGISTRATION NUMBER: 30753
; REFERENCE/DOCKET NUMBER: 7937-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Parvovirus
; US-08-336-345-2

Query Match 51.5%; Score 1040.4; DB 1; Length 5049;
Best Local Similarity 71.4%; Pred. No. 1.2e-292;
Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;

QY 1 ATGGCTGGAAATCGTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTTAAAGGAAAA 60
DB 273 ATGCTGGCAACCACTATCTGAGGAAGTTATGGAGGAGTAATTTGGTTAAGAAACAT 332
QY 61 AGTAACAGGAAGTGTCTCAATTTGTTTAAAAATGAAATGTTCAACTGGAATGGAATA 120
DB 333 GCAGAAAAATGAGCAATTTTGGTTTGTGTTTAAATGTGCAACCTGCAACTTAATGGAAG 392
QY 121 GATATCGGATGGAATAGTTACAAAAGAGCTGCAGGAGGACGAGCTGAAATCTTTACAA 180
DB 393 GATGTTGCTGGAACTAATACCAACCAATTCAAAATGAAGAGCTTAACATCTTTAAT 452
QY 181 CGAGGAGCGAAATCTACTTGGGACCAAGC-----GAGGACATGGAATGGGAAACCA 234
DB 453 AGAGGAGCAACACAGCAATGGATCAACCGAAGAGAGAAATGACCTGGGAATCGGA 512
QY 235 GTGGATGAAATGACCAAAAAGCAAGTATTCATTTTGTTCCTTTGTTTAAAAATGTTA 294
DB 513 GTTGATAGTCTCGCAAAAAGCAAGTACAACTTTTGTGATGCAATTAATTAATAATGCTT 572
QY 295 TTTGAAGTGTTAACCAAGAATATATTTCTCGGTGATGTTAATTTGTTTGTGCAACAT 354
DB 573 TTTGAAGTCTTTGTTCTAAAAATATAGAACCAAAATGATGTTTGGTTTATTCACAT 632
QY 355 GAATGGGAAAAAGCAAGCTGGCACTGCTATCTAATTTGGAGGAAGGACTTTTAGT 414
DB 633 GAATGGGAAAAAGATCAAGCTGGCAATGTCATGTTTACTTCACTAGTAAGAACTTACAA 692
QY 415 CAAGCTCAAGGGAATGTTGGAGAGGCAACTAAATGTTTACTGAGACAGATGGTTGGTA 474
DB 693 CAAGCAACTGGTAAATGGCTACGACAGCAAAATGAATATGATTTATGGAGTAGATGGTTGGTG 752
QY 475 ACAGCCCTGTAATGTGCAACTAACACAGCTGAAAGAAATTAACCTAAGAGAAATAGCAGAA 534

753 ACTCTTTGTTGGTAAACTTAAACCAACTGAAGAGATTAAAGCTCAGAGAAATTCAGAA 812
535 GACAATGAGTGGGTATCTCTACTTATTAAGCAATAAGCAAAACCAAAAGACTATATACC 594
813 GATAGTGAATGGGTGACTATATTAACATACAGACATAAGCAAAACCAAAAGACTATGTT 872
595 AAGTGTGTTCTTTTGGAAACATGATGCTTACTATTTTAACTAAAGAAAGAAATAAGC 654
873 AAAATGGTTCATTTTGGAAATATGATAGCATATATCTTTTAAAGAAAGAAATAATGTC 932
655 ACTAGTCCACCAAGAGACGGAGGCTATTTCTTAGCAGTACTCTGCTCGGAAACATAAC 714
933 CACAT---GACAAAGAAAGTGGCTATTTTAAAGTACTGATCTGCTGGTGGAAATTTAAC 989
715 TTTTAAAGAAAGCGGAGCGCCATCTAGTGAGCAAACTATATACATGATGACATGCGGCCA 774
990 TTTATGAAGTATCAAGACAGACAAAATGTCAGCACACTTTACACTGAACAAATGAACCA 1049
775 GAAACGGTTGAACCACTGTAACCACTGCGCAGGAACTAAGCGCGCAGAAATTCAAACT 834
1050 GAAACCGTTGAACCACTGTAACCACTGCGCAGGAACTAAGCGCGGAGAAATTCAAACT 1109
835 AAAAAAGAAAGTTTCTATTAAACTACACTTAAAGAGCTGGTGCATAAAAGAGTAACCTCA 894
1110 AAAAGCAAGTGTCAATCAATGTACTTTGCGGACTTGGTTAGTAAAGAGTAACATCA 1169
895 CCAGAGACTGGATGATGATGACGACGACAGTGTACATTTGAAATGATGGCTCAACAGGT 954
1170 CTTGAAGACTGGATGATGTTACAAACGATAGTTATATTGAAATGATGGCAACACGGA 1229
955 GGAGAAACCTGCTGAAATAATGCTAGAGATTTGTACACTAATCTTAGCAGAACCAAA 1014
1230 GGTGAAATCTTTTAAATAATACACTTGAATTTGTACTTTGACTTTAGCAGAACCAAA 1289
1015 ACAGCATTTGACTTAATTTAGAAAAAGCTGAAACCAAGCAAACTAAACCACTTTTCACTG 1074
1290 ACNGCATTTGAAATTAATCTTGAAGAGCAGATTAATCTAATCTAATCTTTGATCTT 1349
1075 CTTGACACAAAGAACCTGCAAGATTTTGTCTTTTCTGGCTGGAACTATGTTAAAGTTGC 1134
1350 GCAAAATCTAGAACATGTCAAAATTTTGAATGCAAGGATGGAATTTGATTTAAAGTTGT 1409
1135 CATGCTATTGCTGTGTTTAAACACACAAAGGAGCAAAAGAAATACTGTTTATTTTCAT 1194
1410 CAGCTATATAGCATGTGTTTAAATAGCAAGGTTGTTAAAGAAATACAGTTCTTTTCAT 1469
1195 GGACCAAGCAGCAGCAAGCAAACTCTATTGTCACCAAGCAATAGCAAGCAAGTTGGCAAT 1254
1470 GGACCAAGCAAGTACAGAAATCTATCTGCTCAAGCCATAGCAGCAAGCTGGGTAT 1529
1255 GTTGGTTGCTATAATGACGCAATGTAAACTTTTCCATTTTAAATGACTGTACCAAGAAC 1314
1530 GTTGGTTGTTATAATGACGCAAAATGTAAATTTTCCATTTTAAATGACTGTACCAATAAAT 1589
1315 TTGATTTGGGTAGAAAGCTGGTAACTTTTGGACAGCAAGTAAACAGTTTAAAGCCATT 1374
1590 TTAATTTGGATTCAGAAAGCTGGTAACTTTTGGTCAACAGTTTAAATCAATTTTAAAGCAATC 1649
1375 TGCTCTGCTCAAGCTTATGCTGCTTGAATCAAAAGGAAAGGAGCAAGCAAGATTGAACCA 1434
1650 TGTTCTGGGACAAACAAATTTGATCAAAAGGTAAGGAAAGTAAAGCAATTTGAACCA 1709
1435 ACACCAAGTCAATCATGACCAAAATGAGAACTATTAAGTGGTCTGAGAAATAGGCTCGAAGAA 1494
1710 ACTCCAGTAAATATGACAACTAATGAATAATATAACAAATTTGAGAAATTTGAGAA 1769
1495 AGACCAAGAACACCTCAACTCAACTGACAGACAGAAATGCTTAACTTCACTTAAACATACC 1554
1770 AGACCTGAACATACACAACTCAACTGAGACAGAAATGTTGAACATTAAGTTAGTTAG 1829
1555 TTGCTCTGCTGATGCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1614
1830 CTTCCAGGAGACTTTGGTTGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTA 1889

1615 GTAAAGATGGTTTACCAATCTACCATGGCAAGTCTACTGTCTAAATGGGCAAGTTCTCT 1674
1890 GTTAAACATGGTTTGAATCAACCATGGCTACTATACACATCATTTGGGGAAAGTACCA 1949
1675 GATTGTGTAGAAATCTGGGGGAGCAAGGTGCGCAACTCTTATATAATTTACTAGTTGCG 1734
1950 GAATGGATGAAATCTGGGGGAGCTTAAATACAAAGAGGTATATAATTCACCGGTTGC 2009
1735 GCAGGCTCACCATTACGACACCGAAGTACGCTCTCAGCGAGCAATATGCACTAACT 1794
2010 AAGAGCT---TAGAGACACAAGCGGCAAGCAATCTCAGAGTCAAGACCAAGTTCTTA 2066
1795 CCACTTTGATCGGATCTCGAGGACCTGGCTTTAGAGCCTTTGGAGCACACCAATATCTCT 1854
2067 CTTCTGACTCCGAGCTAGTGGACCTTGCACTGACGACCGTGGAGTACTCCAGATACGCT 2126
1855 GTTGGCGGCACTCGAGAAACCCAGAACCTGTTGGGGAAGTGTTCCTCAAGCTCGCAAGAT 1914
2127 ATTGCAGAACTGCAATCAACCAATCAACCAACTTTGGCGTTACTCA-----CAAAGAC 2180
1915 GGTCAACTGACCCCACTTGGTCAGAGATCGAGAGGATTTGAGAGCGTCTTCGGTGGCG 1974
2181 GTGCAAGCGAGTCCGACGTGGTCCGAAATAGAGCGAGACTGAGAGCCATCTTTACTTCT 2240
1975 GAACCGTTGAAGAAAGACTTTCAGCGA 2000
2241 GAACAATTGGAAGAGATTTTCGAGA 2266

RESULT 3

US-08-647-655-1
; Sequence 1, Application US/08647655
; Patent No. 5885585
; GENERAL INFORMATION:
; APPLICANT: Parrish, Colin R.
; APPLICANT: Gruenberg, Allen
; APPLICANT: Carmichael, Leland E.
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/647,655
; FILING DATE: On Even Date Herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gordon, Jennifer
; REGISTRATION NUMBER: 30,753
; REFERENCE/DOCKET NUMBER: 7937-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Parvovirus

US-08-647-655-1

Query Match 51.5%; Score 1040.4; DB 2; Length 5049;
Best Local Similarity 71.4%; Pred. No. 1.2e-292;
Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;
QY 1 ATGGCTGGAATGCTTACTCTGATCAAGTTTGGGAGCAACCACTGGTTAAAGAA 60
Db 273 ATGTCTGGCAACCACTATATCTGAGGAAGTTATGGAGGAGTAAATTTGGTTAAAGAAACAT 332
QY 61 AGTAACCCAGGAAGTGTCTCTCAATTTCTTTTAAAAATGAAAAATGTTCAACTGAATGAAAA 120
Db 333 GCAGAAATGAAGCAATTTTCGTTTGTAAATGTGCAACGTCCTCACTAATGGAAG 392
QY 121 GATATCGGATGGAATAGTTTACAAAAAGAGTGTGAGGAGGACGAGCTGAAATCTTTACAA 180
Db 393 GATGTTGCGTGGAACAACTATATACCAACCAATTTCAAAATGAAGAGCTAACATCTTTAAAT 452
QY 181 CGAGGAGCGGAACCTACTTGGGACCAAGC-----GAGGACATGGAATGGGAACCA 234
Db 453 AGAGGAGCAACAAACAGCAATGGATCAACCGAAGAAAGAAATGACCTGGGAATCGGA 512
QY 235 GTGGATGAAATGACCAAAAAAGCAAGTATTCTATTTTGTGTTTCTTTGTTAAAAATGTTTA 294
Db 513 GTTGTATGCTCGCCAAAAGCAAGTACAACTTTTGTGATTAATTTAAATAATGCTCT 572
QY 295 TTTGAAGTCTTTAACACAAAGAAATATATTTCTGTTGATGTTAAATTTGTTTGTGCAACAT 354
Db 573 TTTGAAGTCTTTTCTAAAAATATAGAAACCAATGATGTGTTTGGTTTATTCAACAT 632
QY 355 GAATGGGAAAGACCAAGCTGGCACTGCCATGACTAAATTTGGAGGAAGAGCTTTAGT 414
Db 633 GAATGGGAAAGATCAAGGCTGGCAATGTCTGATGTTTCTTCTATAGTAAGAACTTACAA 692
QY 415 CAAGCTCAAGGAAATGGTGGGAAGGCACTAAATGTTTCTTGGAGCAGATGGTTGGTA 474
Db 693 CAAGCACTGTTAATGGCTACGACACAAATGATATGTTTGGAGTAGATGGTTGGTG 752
QY 475 ACAGCTGTAATGTGCAACTAACACAGCTGAAAGAAATTAATAAGAGAAATAGCAGAA 534
Db 753 ACTCTTTGTTGCGTAAACTTAACACCACTGAAAGAGATTAAGCTCAGAGAAATTCAGAA 812
QY 535 GCAATGAGTGGTTACTCTCTACTTATATAGCATAAGCAACCAACCAAGAGCTATACC 594
Db 813 GATAGTAATGGTGACTATATTAACATACAGCATTAAGCAACCAACCAAGAGCTATGTT 872
QY 595 AAGTGTGTTCTTTTGGAAACATGATTGCTTACTATTTTAACTAAAAAGAAATTAAGC 654
Db 873 AAAATGTTCAITTTTGGAAATATGATAGCATATTACTTTTAAACAAAGAAAAATTTGTC 932
QY 655 ACTAGTCCCAAGAGACGAGCGCTATTTTCTTAGCAGTGACTCTGCTGGTGGAAACCTTA 714
Db 933 CACAT---GACAAAAGAAAGTGGCTATTTTAAAGTACTGATTCTCTGGTGGAAATTTAAC 989
QY 715 TTTTAAAGAGAGCGAGCGCCATCTAGTGAGCAAACTATACACTGATGACATCGCGCCA 774
Db 990 TTTATGAAGTATCAAGACAGCAAAATGTCAGCACACTTTTACACTGACAAATGAAACCA 1049
QY 775 GAAACGGTTGAAACCCACAGTAACTCCTGCGCAGGAACTAAGCGCGCAGAAATTCAAACT 834
Db 1050 GAAACCGTTGAAACCCACAGTAACTCCTGCGCAGGAACTAAGCGCGCAGAAATTCAAACT 1109
QY 835 AAAAAAGAGTTTCTATTTAAACTACACTTAAAGAGCTGGTGCATAAAGAGTAACTCA 894
Db 1110 AAAAAAGAGTGTCAATCAATGTAATTTGCGGAGCTTTGCTAGTAAAGAGTAACTCA 1169
QY 895 CCAGAGGACTGGATGATGATGACGACAGAGTGTACATTTGAAATGATGGCTCAACAGGT 954
Db 1170 CTTGAGAGCTGGATGATGATGACCAACAGATAGTTTATTTGAATGATGGCACAACAGGA 1229
QY 955 GGAGAAAACTGCTGAAAAATACGCTAGAGATTTGTACACTAATCTAGCAGAACCAAA 1014
Db 1230 GGTGAAATCTTTTAAAAAATACACTTGAAATTTGTACTTTGACTTTTAGCAAGAACAAA 1289

```
QY 1015 ACAGCATTTGACCTTAATTTTAGAAAAAGCTGAAACGACGAACTAAACCACTTTTCACTG 1074
D 1290 ACAGCATTTGAAATTAATCTTGAATAAGACGATAATACTAACTAACTTTGATCTT 1349
QY 1075 CTTGACACAGAACCTCGCAATTTTTCCTTTTCATGCTGGAACTATGTTAAAGTTTC 1134
D 1350 GCAATTTCTAGAACATGTCATTTTATAGATGACGATGGAATGGAATTAAGTTTGT 1409
QY 1135 CATGCTATTTGCTGTGTTTAAACAGACAGAGGCAAAAGAAATATCTGTTTATTTTCA 1194
D 1410 CACGCTATAGCATGTGTTTAAATAGACAGAGTGGTAAAGAAATACAGTTCTTTTTCAT 1469
QY 1195 GACACGACGACAGCAATCTATTTATTTGCAACAGCCATAGCAGACAGAGTTGGCAAT 1254
D 1470 GGACACGACAGTACAGGAAATCTATCTATGCTCAAGCCATAGCAGCAAGCTGTGGTAA 1529
QY 1255 GTTGGTGTCTATATGACGCAATGTAAACCTTTCCATTTTAACTGCTACCAAGAAC 1314
D 1530 GTTGGTGTCTATATGACGCAATGTAAATTTTCCATTTTAACTGCTACCAATAAAT 1589
QY 1315 TTGATTTGGTGAAGAGCTGTAACTTTGGACAGCAAGTAAACCAAGTTTAAAGCCATT 1374
D 1590 TTAATTTGGATGAGAGCTGTAACTTTGGTCAACAGTTAATCAATTTAAGCAATC 1649
QY 1375 TGCTCTGGTCAAGCTATTTGCTATGATCAAAAAGGAAAGGACGCAAAAGATTTGAACCA 1434
D 1650 TGTCTGGACAAACAAATAGATTTGATCAAAAAGGTAAGGAAATGAACCA 1709
QY 1435 ACACCATCATCATGACCAAAATGAGAACTTACAGTGTGAGATAGCTGCGAGAA 1494
D 1710 ACTCCAGTAAATGACAACTAATGAAATTAACAAATTTGAGAAATTTGATGTAAGAA 1769
QY 1495 AGACCAAGACACACTCAACCAATCAGACAGCAAGTCTTAACTTCACTTAACACATACC 1554
D 1770 AGACCTGACATACACACCAATAGACAGCAAGTGTGAACATTAAGTTAGTATGTAG 1829
QY 1555 TTGCTCTGGTGAATTTGGTTGGTTGACAAAATGAAATGGCCCATGATTTGCTTGGTTG 1614
D 1830 CTTCCAGGAGACTTTGGTTGGTTGATAAAGAAAGTGGCTTTAATATGTGCATGGTTA 1889
QY 1615 GTAAAGATGTTTACCAATCTACATGCGCAAGTCTGTCTAAATGGGCAAGTTCTCT 1674
D 1890 GTTAAACATGTTTGAATCAACCTGCTTAACATACATCATTTGGGGAAGTACCA 1949
QY 1675 GATTGGTCAGAAAACCTGGGCGGAGCAAGGTCCTTAACTTCTTAAATTTACTAGTTCG 1734
D 1950 GAATGGATGAAATCTGGGCGGAGCTTAAATACAGAGAGTATAAATCACCAGTTGC 2009
QY 1735 GCAGCTCACCATTTACGACACCCGAAAGTACGCTCTCAGCAGAACTATGCACTAACT 1794
D 2010 AAAGACT---TAGAGACACAAGCGGCAAGCAATCTCAGAGTCAAGACCAAGTTCTAACT 2066
QY 1795 CCACTTGCATCGATCTGAGGACCTGGCTTTAGAGCTTTGGAGCACACCAATCTCTCT 1854
D 2067 CCTCTGATCTCGGACGTAGTGGACCTTGCCTGGAACCGGAGTACTCCAGATAGCCT 2126
QY 1855 GTTGGGGGACTGACAGAAACCCAGAACACTGGGGAAGCTGTTTCCAAAGCTGCGAAGAT 1914
D 2127 ATTGCGAATCTGCAATCAACATCAACCACTTGGCGTTACTCA-----CAAGAC 2180
QY 1915 GGTCAACTGAGCCCAACTTTGGTCAGAGATCGAGGAGGATTTGAGAGCGTCTCGGTGCG 1974
D 2181 GTCAAGCGAGTCCGAGCTGTCCGAAATAGAGGACGACCTGAGAGCCATCTTTACTTCT 2240
QY 1975 GAACCGTTGAAGAAAGCTTCAGCGA 2000
D 2241 GAACAAATTTGAAGAAGATTTTCGAGA 2266
```

RESULT 4
US-08-647-655-2
; Sequence 2, Application US/08647655

```
; Patent No. 5885585  
; GENERAL INFORMATION:  
; APPLICANT: Parrish, Colin R.  
; APPLICANT: Gruenberg, Allen  
; APPLICANT: Carmichael, Leland E.  
; TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE: US/08/647,655  
; FILING DATE: On Even Date Herewith  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gordon, Jennifer  
; REGISTRATION NUMBER: 30,753  
; REFERENCE/DOCKET NUMBER: 7937-008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864/9741  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5049 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Parvovirus  
; US-08-647-655-2
```

```
Query Match 51.5%; Score 1040.4; DB 2; Length 5049;  
Best Local Similarity 71.4%; Pred. No. 1.2e-292;  
Matches 1432; Conservative 0; Mismatches 556; Indels 18; Gaps 4;  
  
QY 1 ATGGCTGGAAATGCTTACTCTGATGAAGTTTTCGGAGCAACCACTGGTTTAAAGGAAAA 60  
D 273 ATGCTGGCAACCAAGTATCTAGGAGGATTAATGGTTTAAAGAAACAT 332  
  
QY 61 AGTAACGAGGAAGTGTCTCATTTGTTTAAATAATGAAATGTTCAACTGGAATGAAAA 120  
D 333 GCAGAAATGAAGCATTTTCGTTTGTAAATGTGCAACCTCAACTTAAATGGAAAG 392  
  
QY 121 GATATCGGATGGAATAGTTACAAAAAGAGCTCGAGGAGCAGAGCTGAATCTTTACAA 180  
D 393 GATGTTTCGTGGCAACACTATACCAACCAATTTCAAAATGAAGAGCTAACATCTTTAAT 452  
  
QY 181 CGAGGAGCGGAACACTTCTGGGACCAAGC-----GAGGACATGGAATGGGAAACACA 234  
D 453 AGAGAGGACCAACAGCAATGGATCAACCGGAAGAGAGAAATGGAATGGGAATCGGAA 512  
  
QY 235 GTGGATGAAATGACCAAAAAAGCAAGTATTTCAATTTTGTGTTTAAATAATGTTTA 294  
D 513 GTTGATAGTCTCGCAAAAAAGCAAGTACAACTTTTGTGATGATTAATTAATAATGTTT 572  
  
QY 295 TTTGAAGTGTCTTAACCAAAAGAAATATTTCTCGGTGATGTTAATTTGTTTGTGCAACAT 354  
D 573 TTTGAAGTCTTTGTTCTAAAAATATAGAACCAAAATGAATGAATGTTTGGTTTATTCACAT 632  
  
QY 355 GAATGGGAAAGACCAAGCTGGCACTGCCATGTACTAATTTGGAGGAAGGACTTTAGT 414  
D 633 GAATGGGAAAGATCAAGCTGGCATTTGCTTTTACTTTTACTTATAGTAAGAACTTACAA 692
```



```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-254-358-1

Query Match      5.3%; Score 107.8; DB 1; Length 4680;
Best Local Similarity 53.4%; Pred. No. 8.6e-21;
Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

QY 1168 GGCAAAAGAAATACGTGTTTATTTTCATGGACCGCAGCAGCAAGCAATCTATTATTGCA 1227
Db      |||||
QY 1293 GGCAAGAGGAACACCATCTGGCTGTTTGGCCTGCAACTACCGGAAGACCAACATCGCG 1352
Db      |||||
QY 1228 CAAGCCATAGCACAAAGCAGTTGGCAATGTTGGTTGCTTATATGCAAGCAATGTAACATTT 1287
Db      |||||
QY 1353 GAGGCCATAGCCACACATCTGCGCTTCTACGGGTGCGTAAACTGGACCAATGAGAACTTT 1412
Db      |||||
QY 1288 CCATTTAATCACTGTACCAACAAGAACTTGAATTTGGGTAGAAGAGCTGGTAACATTTGGA 1347
Db      |||||
QY 1413 CCCTTCAACGACTGTGTGCAAGATGGTGTCTGTGGGAGGAGGGAAGATGACCGCC 1472
Db      |||||
QY 1348 CAGCAAGTAAACCAAGTTTAAAGCCATTTGCTGCTCAAGCTATTTCGCAATTTGATCAAAA 1407
Db      |||||
QY 1473 AAGTGTGAGTGTGCGCAAGCAATCTCGAGGAGCAAGGTGCGCGTGGACCAAGAA 1532
Db      |||||
QY 1468 ACAGTGTGAGATAGCTCGGAAGAAAGACCAAGCACTCATGACCAACAATGAGAACAT 1527
Db      |||||
QY 1593 TGGCGCGTGTATTGACGGGAACCTCAACGACCTTCGAAACACGAGCGCGTTGCAAGACCG 1652
Db      |||||
QY 1528 ATGCTTAAACATTCATCTAAACACATACCTGCTGCTGGTACATTTGTTGGTTGACAAAAT 1587
Db      |||||
QY 1653 ATGTTCAAAATTTGAATCAACCCCGCTGCTGATCATGACTTTGGGAAGGTTCACCAAGCAG 1712
Db      |||||
QY 1588 GAA 1590
Db      |||||
QY 1713 GAA 1715
Db      |||||
```

```
RESULT 6
US-08-475-391-1
; Sequence 1, Application US/08475391
; Patent No. 5786211
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,391
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/254,358
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5786211and, Greta E.
```

```
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31975
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-475-391-1

Query Match      5.3%; Score 107.8; DB 1; Length 4680;
Best Local Similarity 53.4%; Pred. No. 8.6e-21;
Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

QY 1168 GGCAAAAGAAATACGTGTTTATTTTCATGGACCGCAGCAGCAAGCAATCTATTATTGCA 1227
Db      |||||
QY 1293 GGCAAGAGGAACACCATCTGGCTGTTTGGCCTGCAACTACCGGAAGACCAACATCGCG 1352
Db      |||||
QY 1228 CAAGCCATAGCACAAAGCAGTTGGCAATGTTGGTTGCTTATATGCAAGCAATGTAACATTT 1287
Db      |||||
QY 1353 GAGGCCATAGCCACACATCTGCGCTTCTACGGGTGCGTAAACTGGACCAATGAGAACTTT 1412
Db      |||||
QY 1288 CCATTTAATCACTGTACCAACAAGAACTTGAATTTGGGTAGAAGAGCTGGTAACATTTGGA 1347
Db      |||||
QY 1413 CCCTTCAACGACTGTGTGCAAGATGGTGTCTGTGGGAGGAGGGAAGATGACCGCC 1472
Db      |||||
QY 1348 CAGCAAGTAAACCAAGTTTAAAGCCATTTGCTGCTCAAGCTATTTCGCAATTTGATCAAAA 1407
Db      |||||
QY 1473 AAGTGTGAGTGTGCGCAAGCAATCTCGAGGAGCAAGGTGCGCGTGGACCAAGAA 1532
Db      |||||
QY 1468 ACAGTGTGAGATAGCTCGGAAGAAAGACCAAGCACTCATGACCAACAATGAGAACAT 1457
Db      |||||
QY 1593 TGGCGCGTGTATTGACGGGAACCTCAACGACCTTCGAAACACGAGCGCGTTGCAAGACCG 1592
Db      |||||
QY 1528 ATGCTTAAACATTCATCTAAACACATACCTGCTGCTGGTACATTTGTTGGTTGACAAAAT 1587
Db      |||||
QY 1653 ATGTTCAAAATTTGAATCAACCCCGCTGCTGATCATGACTTTGGGAAGGTTCACCAAGCAG 1712
Db      |||||
QY 1588 GAA 1590
Db      |||||
QY 1713 GAA 1715
Db      |||||

RESULT 7
US-08-709-609-1
; Sequence 1, Application US/08709609
; Patent No. 5858775
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/709,609
/ FILING DATE:
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: No. 5858775and, Greta E.
/ REGISTRATION NUMBER: 35,302
/ REFERENCE/DOCKET NUMBER: 31975
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (312) 474-6300
/ TELEFAX: (312) 474-0448
/ TELEX: 25-3856
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 4680 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ US-08-709-609-1

Query Match 5.3%; Score 107.8; DB 2; Length 4680;
Best Local Similarity 53.4%; Pred. No. 8.6e-21;
Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0;
QY 1168 GGCAGGAGAAATACCTGTTTATTTTCATGGACCCAGCAGCAGCAAAATCTATTATTGCA 1227
DB 1293 GGCAGGAGGACCACTCTGGCTGTTTGGGCTTGCAACTACCGGGAAGACCAACATCGCG 1352
QY 1228 CAAGCCATAGCACAAGAGTTGGCAATGTTGGTTGCTATTAATGCAGCCCAATGTAACATTT 1287
DB 1353 GAGGCCATAGCCACACACTGTGCCCTTCTACGGGTGGCTAACTGGACCAATGAGAACTTT 1412
QY 1288 CCATTTAATGACTGTACCAACAGACTTGAATTTGGGTAGAAAGCTGTTGATTTGGA 1347
DB 1413 CCCTTCAACGACTGTGCGCAAGCACTCTCGGAGGAAGCAAGTGGGAGGAGGGAAGATGACCGCC 1472
QY 1348 CAGCAAGTAAACAGATTGAAACCAACAGTATGATTTGGTCAAGCTATTGCGCAATTTGATCAAAA 1407
DB 1473 AAGTCTGTGAGTGGCCAGATAGACCCGACTCCGCTGATGTCACCTCCACCAACATG 1592
QY 1468 ACAGTGTGAGATAGCTCGGAAGAAAGACCACTCAACCAATCAGAGACAGA 1527
DB 1593 TGGCCGCTGATTGACGGGAACCTCAACGACCTTCAACACAGCAGCGCTTGCAGAGCCGG 1652
QY 1528 ATGCTTAAACATTCATCTAACACATACCTTGCCTGGTGACTTTGGTTTGTGACAAAAT 1587
DB 1653 ATGTTCAAAATTTGAATCAACCCCGCTCTGATCATGACTTTGGGAAGGTTCACCAAGCAG 1712
QY 1588 GAA 1590
DB 1713 GAA 1715

RESULT 8
PCT-US95-07178-1
; Sequence 1, Application PC/TUS9507178
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and Methods
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606

/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US95/07178
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Noland, Greta E.
/ REGISTRATION NUMBER: 35,302
/ REFERENCE/DOCKET NUMBER: 31975
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (312) 474-6300
/ TELEFAX: (312) 474-0448
/ TELEX: 25-3856
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 4680 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ PCT-US95-07178-1

Query Match 5.3%; Score 107.8; DB 5; Length 4680;
Best Local Similarity 53.4%; Pred. No. 8.6e-21;
Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0;
QY 1168 GGCAGGAGAAATACCTGTTTATTTTCATGGACCCAGCAGCAGCAAAATCTATTATTGCA 1227
DB 1293 GGCAGGAGGAAACCACTCTGGCTGTTTGGGCTTGCAACTACCGGGAAGACCAACATCGCG 1352
QY 1228 CAAGCCATAGCACAAGCAGTTCGCAATGTTGGTTGCTATTAATGCAGCCCAATGTAACATTT 1287
DB 1353 GAGGCCATAGCCACACACTGTGCCCTTCTACGGGTGGCTAACTGGACCAATGAGAACTTT 1412
QY 1288 CCATTTAATGACTGTACCAACAGACTTGAATTTGGGTAGAAAGCTGTTGATTTGGA 1347
DB 1413 CCCTTCAACGACTGTGCGCAAGCACTCTCGGAGGAGGGAAGATGACCGCC 1472
QY 1348 CAGCAAGTAAACAGATTGAAACCAACAGTATGATTTGGTCAAGCTATTGCGCAATTTGATCAAAA 1407
DB 1473 AAGTCTGTGAGTGGCCAGATAGACCCGACTCCGCTGATGTCACCTCCACCAACATG 1532
QY 1408 GGAAGAGGAGCAACAGATTTGAACCAACAGTCTCATGACCAACAAATGAGAACAT 1467
DB 1533 TGCAGTCTCGGCCAGATAGACCCGACTCCGCTGATGTCACCTCCACCAACATG 1592
QY 1468 ACAGTGTGAGATAGCTCGGAAGAAAGACCACTCAACCAATCAGAGACAGA 1527
DB 1593 TGGCCGCTGATTGACGGGAACCTCAACGACCTTCAACACAGCAGCGCTTGCAGAGCCGG 1652
QY 1528 ATGCTTAAACATTCATCTAACACATACCTTGCCTGGTGACTTTGGTTTGTGACAAAAT 1587
DB 1653 ATGTTCAAAATTTGAATCAACCCCGCTCTGATCATGACTTTGGGAAGGTTCACCAAGCAG 1712
QY 1588 GAA 1590
DB 1713 GAA 1715

RESULT 9
US-09-807-802A-18
; Sequence 18, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNVN.031USA

; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 18
; TYPE: DNA
; ORGANISM: AAV-2
; US-09-807-802A-18

Query Match 5.3%; Score 107.8; DB 4; Length 4681;
Best Local Similarity 53.4%; Pred. No. 8.6e-21;
Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0;
QY 1168 GGCAGAGGAACTCTGTTTATTTTCATGGACCGCCAGCACAGGCAAAATCTATTATGCA 1227
DB 1293 GGCAGAGGAACTCTGTTTATTTTCATGGACCGCCAGCACAGGCAAAATCTATTATGCA 1352
QY 1228 CAAAGCCATAGCACCAAGCAGTTGGCAATGTTGGTTGCTATTAATGCAGCCAAATGTAACATTT 1287
DB 1353 GAGCCATAGCCCACTGTGCGCTTCTACGGGTGCTAACTGGACCAATGAGAACTTT 1412
QY 1288 CCATTAAATGACTGTACCAAGAACTTGTATTGGGTAGAGAAAGCTGGTAACTTTGGA 1347
DB 1413 CCCTTCAACGACTGTGTCACAGATGGTGTATCTGTGGGAGGAGGAGATGACCGCC 1472
QY 1348 CAGCAAGTAAACGAGTTAAAGCCATTTGCTGTGTCAGCTATTTCGATTTGATCAAAAA 1407
DB 1473 AAGTCGTGAGTCCGCGCAAGCAATCTCGAGGAAGCAAGTGGCGGTGGACCAAGAA 1532
QY 1408 GGAAGGAGGAGCAAAAGATTTGAACCAACCAAGTATCATATGACCAATGAGACATT 1467
DB 1533 TGCAAGTCTTCGCGCCAGATAGACCCGACTCCCGTGTATGTCACCTCCCAACCAATG 1592
QY 1468 ACAGTGTGTAGATAGCTCGGAAGAAAGACAGAAACACACTCAACCAATCAGAGACAGA 1527
DB 1593 TGCGCGTGTATTCAGGGAACTCAACGACCTTCGAACACAGCAGCGGTTGCAAGCCGG 1652
QY 1528 ATGCTTAAACATTCATCAACATACCTTCGCTGGTGTGATCTTGGTTGGTTGACAAAAAT 1587
DB 1653 ATGTTCAAATTTGAATCACTACCCCGCTGATCATGACTTTGGGAAGGTTCACCAAGCAG 1712
QY 1588 GAA 1590
DB 1713 GAA 1715

RESULT 10
US-08-331-384-2/c
; Sequence 2, Application US/08331384
; Patent No. 5856152
; GENERAL INFORMATION:
; APPLICANT: Wilton, James M.
; APPLICANT: Kelley, William M.
; APPLICANT: Fisher, Krishna J.
; TITLE OF INVENTION: Hybrid Adenovirus-AAV Vector and Methods
; TITLE OF INVENTION: of Use Thereof
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P.O. Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/331,384
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: UPNG1149USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4910 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; US-08-331-384-2

Query Match 5.3%; Score 107.8; DB 2; Length 4910;
Best Local Similarity 53.4%; Pred. No. 8.9e-21;
Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0;
QY 1168 GGCAGAGGAACTCTGTTTATTTTCATGGACCGCCAGCACAGGCAAAATCTATTATGCA 1227
DB 1544 GGCAGAGGAACTCTGTTTATTTTCATGGACCGCCAGCACAGGCAAAATCTATTATGCA 1485
QY 1228 CAAAGCCATAGCACCAAGCAGTTGGCAATGTTGGTTGCTATTAATGCAGCCAAATGTAACATTT 1287
DB 1484 GAGCCATAGCCCACTGTGCGCTTCTACGGGTGCTAACTGGACCAATGAGAACTTT 1425
QY 1288 CCATTAAATGACTGTACCAAGAACTTGTATTGGGTAGAGAAAGCTGGTAACTTTGGA 1347
DB 1424 CCCTTCAACGACTGTGTCACAGATGGTGTATCTGTGGGAGGAGGAGATGACCGCC 1365
QY 1348 CAGCAAGTAAACGAGTTAAAGCCATTTGCTGTGTCAGCTATTTCGATTTGATCAAAAA 1407
DB 1364 AAGTCGTGAGTCCGCGCAAGCAATCTCGAGGAAGCAAGTGGCGGTGGACCAAGAA 1305
QY 1408 GGAAGGAGGAGCAAAAGATTTGAACCAACCAAGTATCATATGACCAATGAGACATT 1467
DB 1304 TGCAAGTCTTCGCGCCAGATAGACCCGACTCCCGTGTATGTCACCTCCCAACCAATG 1245
QY 1468 ACAGTGTGTAGATAGCTCGGAAGAAAGACAGAAACACACTCAACCAATCAGAGACAGA 1527
DB 1244 TGCGCGTGTATTCAGGGAACTCAACGACCTTCGAACACAGCAGCGGTTGCAAGCCGG 1185
QY 1528 ATGCTTAAACATTCATCAACATACCTTCGCTGGTGTGATCTTGGTTGGTTGACAAAAAT 1587
DB 1184 ATGTTCAAATTTGAATCACTACCCCGCTGATCATGACTTTGGGAAGGTTCACCAAGCAG 1125
QY 1588 GAA 1590
DB 1124 GAA 1122

RESULT 11
US-08-836-087-2/c
; Sequence 2, Application US/08836087
; Patent No. 5871982
; GENERAL INFORMATION:
; APPLICANT: Trustees of University of Pennsylvania
; APPLICANT: Wilton, James M.
; APPLICANT: Kelley, William M.
; APPLICANT: Fisher, Krishna J.
; TITLE OF INVENTION: Hybrid Adenovirus-AAV Vector and
; TITLE OF INVENTION: Methods of Use Thereof
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, PO Box 457

```
/ CITY: Spring House
/ STATE: Pennsylvania
/ COUNTRY: USA
/ ZIP: 19477
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/836,087
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/331,384
/ FILING DATE: 28-OCT-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Bak, Mary E.
/ REGISTRATION NUMBER: 31,215
/ REFERENCE/DOCKET NUMBER: GNVPN.007PCT
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 215-540-9200
/ TELEFAX: 215-540-5818
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 4910 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: unknown
/ MOLECULE TYPE: cdna
/ US-08-836-087-2

Query Match 5.3%; Score 107.8; DB 2; Length 4910;
Best Local Similarity 53.4%; Pred. No. 8.9e-21;
Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

QY 1168 GGCAGAAAGAAATCTGTTTATTTTCATGGACCGCAGCAGCAGCAATCTATTATGCA 1227
DB 1544 GGCAGAGGAACACCATCTGGCTGTTTGGGCTTGCCTGCACTACCGGGAAGACCAATCGCG 1485
QY 1228 CAAGCCATAGCACAAGCAGTTGGCAATGTTGGTGTCTATTAATGCAGCCAAATGTAACATTT 1287
DB 1484 GAGGCCATAGCCCACTGTGCCCTTCTAGGGTGGTAACTGGGACCAATGAGAATTT 1425
QY 1288 CCATTTAATGACTGTACCAACAGAACTTTGATTGGGTAGAAGAAGCTGGTAACTTTGGA 1347
DB 1424 CCCTTCAACGACTGTCTGCACAAGATGGTGATCTGTGGGAGGAGGAGATGACCGCC 1365
QY 1348 CAGCAAGTAAACAGTTTAAAGCCATTTCCTGGTCAAGCTATTTCGCTTGTATCAAAAA 1407
DB 1364 AAGTCTGTGAGTCGCGCAAGCCATTCTCGGAGGAAGCAAGGTGCGCGTGGACCAAGAA 1305
QY 1408 GGAAGAGGAGCAACAGATTGAACCAACACAGTCTCATGACCAACCAATGAGAATTT 1467
DB 1304 TGCAAGTCTTGGCCCGAGATAGACCGACTCCCGTGATCTGCTACCTCCACCAACATG 1245
QY 1468 ACAGTGGTCAAGATAGCTGCGAAGAAAGACCAAGAACACACTCAACCAATCAGAGACAGA 1527
DB 1184 ATGTTCAAAATTTGAACCTCACCCCGCTGGATCATGACTTTGGGAAGGTCAACCAAGCAG 1125
QY 1588 GAA 1590
DB 1124 GAA 1122

RESULT 12
US-09-246-320-2/c
; Sequence 2, Application US/09246320
; Patent No. 6251677
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Trustees of University of Pennsylvania
/ APPLICANT: Wilson, James M.
/ APPLICANT: Kelley, William M.
/ APPLICANT: Fisher, Krishna J.
/ TITLE OF INVENTION: Hybrid Adenovirus-AAV Vector and
/ TITLE OF INVENTION: Methods of Use Thereof
/ NUMBER OF SEQUENCES: 2
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Howson and Howson
/ STREET: Spring House Corporate Cntr, PO Box 457
/ CITY: Spring House
/ STATE: Pennsylvania
/ COUNTRY: USA
/ ZIP: 19477
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/246,320
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/836,087
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Bak, Mary E.
/ REGISTRATION NUMBER: 31,215
/ REFERENCE/DOCKET NUMBER: GNVPN.007PCT
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 215-540-9200
/ TELEFAX: 215-540-5818
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 4910 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: unknown
/ MOLECULE TYPE: cdna
/ US-09-246-320-2
```

```
Query Match 5.3%; Score 107.8; DB 3; Length 4910;
Best Local Similarity 53.4%; Pred. No. 8.9e-21;
Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

QY 1168 GGCAGAAAGAAATCTGTTTATTTTCATGGACCGCAGCAGCAGCAATCTATTATGCA 1227
DB 1544 GGCAGAGGAACACCATCTGGCTGTTTGGGCTTGCCTGCACTACCGGGAAGACCAATCGCG 1485
QY 1228 CAAGCCATAGCACAAGCAGTTGGCAATGTTGGTGTCTATTAATGCAGCCAAATGTAACATTT 1287
DB 1484 GAGGCCATAGCCCACTGTGCCCTTCTAGGGTGGTAACTGGGACCAATGAGAATTT 1425
QY 1288 CCATTTAATGACTGTACCAACAGAACTTTGATTGGGTAGAAGAAGCTGGTAACTTTGGA 1347
DB 1424 CCCTTCAACGACTGTCTGCACAAGATGGTGATCTGTGGGAGGAGGAGATGACCGCC 1365
QY 1348 CAGCAAGTAAACAGTTTAAAGCCATTTCCTGGTCAAGCTATTTCGCTTGTATCAAAAA 1407
DB 1364 AAGTCTGTGAGTCGCGCAAGCCATTCTCGGAGGAAGCAAGGTGCGCGTGGACCAAGAA 1305
QY 1408 GGAAGAGGAGCAACAGATTGAACCAACACAGTCTCATGACCAACCAATGAGAATTT 1467
DB 1304 TGCAAGTCTTGGCCCGAGATAGACCGACTCCCGTGATCTGCTACCTCCACCAACATG 1245
QY 1468 ACAGTGGTCAAGATAGCTGCGAAGAAAGACCAAGAACACACTCAACCAATCAGAGACAGA 1527
DB 1244 TGCGCGTGTATTGACGGGAACCTCAACGACCTTCCGAACACAGCAGCGGTGCAAGACCGG 1185
QY 1528 ATGCTTAAACATTCATCTAACACATACCTTGCCTGGTGGACTTTGGTTGGTGAACAAAT 1587
DB 1528 ATGCTTAAACATTCATCTAACACATACCTTGCCTGGTGGACTTTGGTTGGTGAACAAAT 1587
```


Db 3539 ATGTTCAAATTTGAATCACC CGCGTCTGGATCATGACTTTGGGAAGGTCACCAAGCAG 3480
QY 1588 GAA 1590
Db 3479 GAA 3477

RESULT 15
US-09-770-315-3
; Sequence 3, Application US/09770315
; Patent No. 6429001
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; TITLE OF INVENTION: Recombinant AAV Packaging Systems
; FILE REFERENCE: 20263-501
; CURRENT APPLICATION NUMBER: US/09/770,315
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,536
; PRIOR FILING DATE: 2000-01-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 7557
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: recombinant DNA
US-09-770-315-3

Query Match 5.3%; Score 107.8; DB 4; Length 7557;
Best Local Similarity 53.4%; Pred. No. 1.1e-20;
Matches 226; Conservative 0; Mismatches 197; Indels 0; Gaps 0;
QY 1168 GGCAGAAAGAAATACTGTTTATTTTCATGGACCGCCAGCAGCGCAATCTATTTATGCA 1227
Db 1321 GGCAGAGGAACACCATCTGGCTGTTTGGGCTTGCACTACCGGGAAGACCAACATCGCG 1380
QY 1228 CAAGCCATAGCACAGCAGTTGGCAATGTTGGTGTCTATTAATGCAGCCAAATGTAACCTTT 1287
Db 1381 GAGGCCATAGCCACACTGTCCTTCTAGGGTGCCTAACTGGACCAATGAGAACTTT 1440
QY 1288 CCATTTAATGACTGTACCAAGAACTTGAATTTGGGTAGAAGAAGCTGGTAACTTTTGA 1347
Db 1441 CCCTTCAACGACTGTGCGACAAGATGGTGTCTGGTGGAGGAGGGAAGATGACCGCC 1500
QY 1348 CAGCAAGTAAACGAGTTTAAAGCCATTTGCTCTGGTCAAGCTATTCGCATTGATCAAAA 1407
Db 1501 AAGGTCGTGGAGTCGGCCAAAGCCATTCTCGGAGGAAGCAAGGTGCGCGTGGACCAAGAA 1560
QY 1408 GGAAGAGGAGCAAAACAGATTGAACCAACACCAAGTCAATCATGACCAAAATGAGAACATT 1467
Db 1561 TGAAGTCCTCGGCGCCAGATAGACCCGACTCCCGTGTATCTGATCTCAACACCAACATG 1620
QY 1468 ACAGTGGTCAAGATPAGGCTCGAAGAAAGACCAAGACACACTCAACCAATCAGAGACAGA 1527
Db 1621 TGGCGCGTGTATTGACGGGAACCTCAACGACCTTGAACACACAGCAGCGCTTGAAGACCGG 1680
QY 1528 ATGCTTAACATTATCTAAACACATACCTTGCCTGGTGCATTTGGTTGGTTCAGAAAAT 1587
Db 1681 ATGTTCAAATTTGAATCACC CGCGTCTGGATCATGACTTTGGGAAGGTCACCAAGCAG 1740
QY 1588 GAA 1590
Db 1741 GAA 1743

Search completed: January 19, 2005, 17:06:01
Job time : 169.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2005, 11:50:09 ; Search time 1072 Seconds

(without alignments)
10821.780 Million cell updates/sec

Title: US-10-069-056-16

Perfect score: 2019

Sequence: 1 atggctggaagtcttactc.....agccgctgaactggactaa 2019

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1771	87.7	5121	18	US-10-647-111A-1
2	941.4	46.6	5075	17	US-10-361-002-3
3	941.4	46.6	5075	17	US-10-361-004-3
4	307.4	15.2	374	18	US-10-466-894-1276
5	285.8	14.2	421	18	US-10-466-894-430
6	284.8	14.1	423	18	US-10-466-894-375
7	284.8	14.1	464	18	US-10-466-894-370
8	284.8	14.1	473	18	US-10-466-894-367
9	284.8	14.1	486	18	US-10-466-894-358
10	284.8	14.1	491	18	US-10-466-894-428
11	284.2	14.1	420	18	US-10-466-894-438
12	283.8	14.1	343	18	US-10-466-894-388

13	283.8	14.1	343	18	US-10-466-894-390	Sequence 390, App
14	283.8	14.1	403	18	US-10-466-894-433	Sequence 433, App
15	283.2	14.0	497	18	US-10-466-894-369	Sequence 369, App
16	280.4	13.9	340	18	US-10-466-894-383	Sequence 383, App
17	278.2	13.8	342	18	US-10-466-894-416	Sequence 416, App
18	273.8	13.6	420	18	US-10-466-894-432	Sequence 432, App
19	269.6	13.4	324	18	US-10-466-894-391	Sequence 391, App
20	263.4	13.0	339	18	US-10-466-894-1275	Sequence 1275, App
21	261.6	13.0	424	18	US-10-466-894-429	Sequence 429, App
22	163.8	8.1	468	18	US-10-466-894-524	Sequence 524, App
23	162.2	8.0	451	18	US-10-466-894-519	Sequence 519, App
24	162.2	8.0	457	18	US-10-466-894-518	Sequence 518, App
25	162.2	8.0	465	18	US-10-466-894-525	Sequence 525, App
26	162.2	8.0	469	18	US-10-466-894-520	Sequence 520, App
27	162.2	8.0	472	18	US-10-466-894-531	Sequence 531, App
28	162.2	8.0	515	18	US-10-466-894-444	Sequence 444, App
29	162.2	8.0	516	18	US-10-466-894-445	Sequence 445, App
30	162.2	8.0	516	18	US-10-466-894-451	Sequence 451, App
31	162.2	8.0	516	18	US-10-466-894-458	Sequence 458, App
32	162.2	8.0	517	18	US-10-466-894-462	Sequence 462, App
33	162.2	8.0	530	18	US-10-466-894-442	Sequence 442, App
34	162.2	8.0	530	18	US-10-466-894-450	Sequence 450, App
35	161	8.0	516	18	US-10-466-894-447	Sequence 447, App
36	160.6	8.0	530	18	US-10-466-894-456	Sequence 456, App
37	159	7.9	468	18	US-10-466-894-523	Sequence 523, App
38	133.4	6.6	247	18	US-10-466-894-471	Sequence 471, App
39	117	5.8	1386	9	US-09-792-630-32	Sequence 32, Appl
40	117	5.8	1386	10	US-09-953-351-32	Sequence 32, Appl
41	117	5.8	1386	13	US-10-080-376-32	Sequence 32, Appl
42	117	5.8	1386	14	US-10-082-671-38	Sequence 38, Appl
43	117	5.8	1386	14	US-10-097-100-32	Sequence 32, Appl
44	117	5.8	1386	15	US-10-023-208-32	Sequence 32, Appl
45	115.4	5.7	1884	9	US-09-792-630-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1

US-10-647-111A-1
; Sequence 1, Application US/10647111A
; Publication No. US20040209240A1
; GENERAL INFORMATION:
; APPLICANT: IGGO, RICHARD
; APPLICANT: VALERBA, MADDALENA
; TITLE OF INVENTION: ANTI-NEOPLASTIC VIRUSES
; FILE REFERENCE: 604-691
; CURRENT APPLICATION NUMBER: US/10/647,111A
; CURRENT FILING DATE: 2003-08-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 1
; LENGTH: 5121
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Parvovirus H1
; OTHER INFORMATION: with promoter P4 and left hairpin from WMV (pph1)
US-10-647-111A-1

Query Match	87.7%	Score	1771	DB	18	Length	5121
Best Local Similarity	92.3%	Pred.	No. 0				
Matches	1864	Conservative	0	Mismatches	155	Indels	0
Gaps	0						
QY	1	ATGCGTGGAAATCCTTACTCTGATGAAGTTTCGGAGCAACCACTGTTAAAGGAAAA	60				
Db	282	ATGCGTGGAAATCCTTACTCTGATGAAGTTTCGGAGCAACCACTGTTAAAGGAAAA	341				
QY	61	AGTAACAGGAGTGTCTCATTTGTTTTTAAATAATGAAATGTTCAACTGGAATGGAATA	120				
Db	342	AGTAACAGGAGTGTCTCATTTGTTTTTAAATAATGAAATGTTCAACTGGAATGGAATA	401				
QY	121	GATATCGGATGGAAATAGTTACAAAAAAGAGCTCGAGGAGCGAGCTGGAATCTTTACAA	180				

;/ CURRENT FILING DATE: 2003-02-10
;/ NUMBER OF SEQ ID NOS: 99
;/ SOFTWARE: PatentIn version 3.2
;/ SEQ ID NO 3
;/ LENGTH: 5075
;/ TYPE: DNA
;/ ORGANISM: Porcine parvovirus
US-10-361-004-3

Query Match 46.6%; Score 941.4; DB 17; Length 5075;
Best Local Similarity 70.4%; Pred. No. 5.7e-237;
Matches 1305; Conservative 0; Mismatches 536; Indels 12; Gaps 3;

```
QY 4 GCTGGAATGCTTACTCTGATGAAGTTTGGGAGCAACCACTGGTTAAAGGAAAAAGT 63
DB 298 GCGGGAACACATTACTCGGAAGAGGTACTAAAGCTACCACTGGCTTCAAGATAATGCT 357
QY 64 AACGAGAGTGTCTTCATTTGTTTAAATGAAATGTTTCAACTGTAATGGAAGAT 123
DB 358 CAAAAAGAGCAATCTCTTATGTATTTAAACACAAAAAGTCAATCTAAATGGAAGAA 417
QY 124 ATCGGATGGAATAGTTACAAAAAGAGCTGCAGGAGGAGCTGAAATCTTTTACAAAGA 183
DB 418 ATTGCTTGAATAACTACAAAGAGATACACAGATGCGGAATGATTAACCTTACAAGA 477
QY 184 GGAGCGAAACTACTTGGGACCAAGCGAGGACATGGAATGGGAAACCACTGAGTGATGAA 243
DB 478 GGAGCAGAAACATCATGGGACCAAGGCAACAGACATGGAATGGGAATCAGAAATCGACAGC 537
QY 244 ATGACCAAAAGCAAGATTCATTTTGTATCTTTGGTTTAAATGTTTAAAGTTG 303
DB 538 CTCACAAAACGCGAAGTACTGATTTTGTACTCTCTTTGTTTAAATGTTCTCTTTGAAGGT 597
QY 304 CTTTACACAAAGATATATTTCTGTGTGATTTAAATGTTTGTCAACATGAATCGGGA 363
DB 598 ATATTGCAAAAGAACCTTAAGTCAAGTACTGCTGTTTCAACAGCATGAATGATGTT 657
QY 364 AAAGACCAAGGCTGGCACTGCCATGTACTAATTTGGAGGAAGGACTTTAGTCAAGCTCAA 423
DB 658 CAAGATATGCGCTATCAC TGCCATGTACTACTAGGTGGAAAGGCTTACAAACAAAGCAATG 717
QY 424 GGAATTTGGGAGAGGCACTAATTTGTTTACTTGAGCAGATGTTGGTAAACAGCTGT 483
DB 718 GGAATTTGGTTCGAAACCAATTTAAACAAATTTTGGATGATGTTTAAATGTTCAATG 777
QY 484 AATGTGCAACTAACACAGCTGAAAGAAATTAAGAAATAGCAGAAAGCAATGAG 543
DB 778 AAAGTACCTTAACACAGTTGAAAGATTAATTAAGGAATAGCAGAGGATGTTGAG 837
QY 544 TGGGTTACTCTTACTTATAAGCATAAGCAACCAAAAGACTATACCAAGTGTGT 603
DB 838 TGGGTTATCGCTACTTAACCTACACTCAAAACAAACTAAAAAACAATATACAAAAATGACT 897
QY 604 CTTTTTGGAAACAATGATGTTTACTATTTTAACTAAAAAGAAATAAGCACTAGTCCA 663
DB 898 CATTITGGAAATATGATGTTCTTACTTCTTAATTAATAAAGAAAGAAAGCAACT----- 951
QY 664 CCAAGAGACGGAGCTATTTCTTAGCAGTGACTCTGGCTGGAAAACTAACTTTTAAAAA 723
DB 952 GAAAGAGAGCATGGATATATCTCAGCTCAGATTTGGCTTCAATGCAAAATTTCTTAAAA 1011
QY 724 GAAGGGAGCGCCATCTAGTGAGCAACCTATACACTGATGACATGCGGCGAGAAACGGTT 783
DB 1012 GAAGCGGAGAGACACTTAGTCAGTCACCTATTTTACTGAACCAATAAACCCTGAAACTGTG 1071
QY 784 GAAACCAAGTAACCACTGGCGAGGAACTAAGCGGCGGAGAAATTAACATAAAGAA 843
DB 1072 GAAACCAACGGTTACTACAGCTCAGGAGCCAAAGAGGAGAGATACAAACAAAAAAGRA 1131
QY 844 GTTCTTATTAACACTACACTTAAAGAGCTGCTGCATAAAGAGTAACTCACCAGAGGAC 903
DB 1132 GTAAGCAATAAATGCAATAAGAGACTGTTGTTAATAAAGATGTACTAGCATAGAAGAC 1191
```

RESULT 4

US-10-466-894-1276
; Sequence 1276, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Anson, Robert
; APPLICANT: Tuijnder, Marius

```
QY 904 TGGATGATGATCGAGCCAGACAGAGTTACATTGAAATGATGGCTCAACACGAGTGGAGAAAC 963
DB 1192 TGGATGATGATCGAGATCCAGACAGTTATATAGAAATGATGGCTCAACACGAGGAGAGAAAT 1251
QY 964 CTGCTGAAAAATACGCTAGAGATTTGTACACTAACTCTAGCCAGAAACCAAAACAGCATTT 1023
DB 1252 TTAATCAAAATATACACTAGAAATATACACTCTTACTCTAGCAAGAAACAAAAACAGCATAT 1311
QY 1024 GACTTAAATTTAGAAAAAGCTGAAACCAAGCAACTAAACCAACTTTTCACTGGCTGACACA 1083
DB 1312 GACTTAAATTTAGAAAAAGCAAAACCAAGCATGCTACCAACATTTTAAATATTAGCAATACA 1371
QY 1084 AGAACCTGAGAAATTTTCTCTTTTCTGCTGGAATATGTTAAAGTTTGGCCATGCTATT 1143
DB 1372 AGAACATGTAAATATTTCAAGCATGCAATTTGGAACCTACATTTAAAGCTTGGCCATGTATA 1431
QY 1144 TCGTGTGTTTAAACAGACAGAGGAGCAAAAGAAATACATGTTTATTTTCTATGGACAGCC 1203
DB 1432 ACTTGTGTACTAAACAGACAGAGGAGAAAGAAATACAAATTTCTATTTTATGGGCCAGCA 1491
QY 1204 AGCAGGCAAAATCTATTATTGCAAGCCATAGCA CAAGCAGTTGGCAATGTTGTTGC 1263
DB 1492 TCAACAGGAAAGATATATTGCTCAACACATTTGCAAACTTAGTTGGTAAATGTTGTTGC 1551
QY 1264 TATAATGCAAGCAATGTAACCTTTTCCATTTTATGACTGTACCAACAGAACTGTTGTTGC 1323
DB 1552 TACAATGCAAGCAATGTAACCTTTTCCATTTTAAATGACTGTACAAATAAAAAATTAATG 1611
QY 1324 GTAGAGAGAGCTGTTAACTTTTGACAGCAAGTAAACCAAGTTTAAAGCCATTTGCTCTG 1383
DB 1612 ATTGAAGAGAGCAGAACTTCTTAACCAAGTAAACCAATTTCAAGCCATATGTTCAAGGT 1671
QY 1384 CAAGCTTATTTCGCAATGATCAAAAGAGGAGCAAAACAGATTGAAACCAACACCAAGTC 1443
DB 1672 CAACAAATTAGAATTTGACCAAAAGGTAAGGAAGCAAAACAAATTTGAACCAACTCTGTA 1731
QY 1444 ATCATGACCAAAATGAGAAACATTACAGTGGTGCAGATAGGCTGCGAAGAGAACCAAGAA 1503
DB 1732 ATAATGACTTACAAATGAAGACATAACTAAAGTTAGAAATAGGATGCGAGAGAAAGACAGAA 1791
QY 1504 CACACTCAACCAATCAGACAGAGAAATGCTTAACATTTCATCAACATACCTTGCCTGCT 1563
DB 1792 CATACAAACCAATAGACAGAGAAATGTTTAAACATAAACCTTAACCAAGAAACCTGCCAGGT 1851
QY 1564 GACTTTGGTTTGGTTGACAAAAATGAATGGCCCATGATTTGCTGTTGGTTGTTGTTGAAAT 1623
DB 1852 GATTTTGGACTTTTAGAAGAACTGAATGGCCCACTAATATGCTTGGTTGGTAAAGAA 1911
QY 1624 GGTACCAATCTACCATGCAAGCTACTGCTTAATGGGCAAAAGTTCTGATTTGGTCA 1683
DB 1912 GGTACCAAGCAACATGGCTAGCTATATGTCATCATTTGGGAAATGTACTGATTTGGTCA 1971
QY 1684 GAAACTGGGCGAGCAAAAGGTGCAACTCTTATAAAATTTACTAGTTTCGGCAGCTCA 1743
DB 1972 GAAATGGAGAGGCAAAATATGCAACCCCAATTAATACACCA---ACAGACTCTCAG 2028
QY 1744 CCATTCACGACACCGAAAAAGTAGCCTCTCAGCCAGAACTATGCACTAATCCACTTGCAT 1803
DB 2029 ATTTCCACATCAGTGAAGAACTTGGCCAGCGGCAACAACTACGCGCAACTCCCAATACAG 2088
QY 1804 TGGATCTCGA---GGACCTGGCTTTAGAGCTTTGGAGCAGCACCAATATCTCC 1853
DB 2089 GAGGACCTGGATTTAGCTTTAGCTTTGGAGCCGTGGAGCGGAGCCAAACCAACACC 2141
```

```
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: US/10/466,894
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; QUERY MATCH: 15.2%; Score 307.4; DB 18; Length 374;
; BEST LOCAL SIMILARITY 88.2%; Pred. No. 2.3e-70;
; Mismatches 0; Indels 0; Gaps 0;
; TYPE: DNA
; ORGANISM: Homo Sapiens
; NAME/KEY: misc feature
; LOCATION: 291_300, 310, 316, 347, 348, 370, 371, 374
; OTHER INFORMATION: n = A,T,C or G
; US-10-466-894-1276

Query Match 15.2%; Score 307.4; DB 18; Length 374;
Best Local Similarity 88.2%; Pred. No. 2.3e-70;
Matches 329; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 174 TTTACACGAGGCGGAACTACTTGGGACCAAGCGAGGACATGGAATGGGAAACAC 233
Db 1 TTTACACGAGGCGGCGAGACCACTTGGGACCAAGCGAGGACATGGAATGGGAGCGC 60

Qy 234 AGTGGATGAATGACCAAAAGCAAGTATTCATTTTGGTCTCTTGGTTAAAAATGTTT 293
Db 61 AGTGGATGACATGACCAAAAGCAAGTATTCATTTTGGTCTCTTGGTTAAAAATGTTT 120

Qy 294 ATTGGAAGTCTTAACCAAGCAATATATTTCTGGTGTATGTTTAAATGTTTGTGCAACA 353
Db 121 GTTTGAAGTCTGACCAAAAGCAATATATTTCTGGTGTATGTTTAAATGTTTGTGCAACA 180

Qy 354 TGAATGGGAAAAGACCAAGCTGGCACTGCCATGTACTAATTTGGAGGAAAGGACTTTAG 413
Db 181 TGAATGGGAAAAGACCAAGCTGGCACTGCCATGTACTAATTTGGAGGAAAGGACTTTAG 240

Qy 414 TCAAGCTCAAGGAAATGGTGGAGAGCAACTAAATGTTTACTGGAGCAGATGGTTGGT 473
Db 241 TCAAGCTCAAGGAAATGGTGGAGAGCAACTAAATGTTTACTGGAGTANATGGTTGGN 300

Qy 474 AACAGCTGTAACTGCAACTTAACCAAGCAAGTATTAATTAAGAGAAATAGCAGA 533
Db 301 GACTGCTGNAATGTNCAACTTAACCAAGCAAGTATTAATTAAGAGAAATAGCAGA 360

Qy 534 AGACAATGAGTGG 546
Db 361 GGACAGTGANNNG 373

RESULT 5
US-10-466-894-430
; Sequence 430, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION...
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: US/10/466,894
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
```

```
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 430
; LENGTH: 421
; TYPE: DNA
; ORGANISM: Homo Sapiens
; NAME/KEY: misc feature
; LOCATION: 276_405, 417, 419, 421
; OTHER INFORMATION: n = A,T,C or G
; US-10-466-894-430

Query Match 14.2%; Score 285.8; DB 18; Length 421;
Best Local Similarity 87.6%; Pred. No. 1.2e-64;
Matches 311; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 173 CTTTACACGAGGCGGAACTACTTGGGACCAAGCGAGGACATGGAATGGGAAACCA 232
Db 59 CTTTACACGAGGCGGCGAGACCACTTGGGACCAAGCGAGGACATGGAATGGGAGAGCG 118

Qy 233 CAGTGGATGAATGACCAAAAGCAAGTATTCATTTTGGTCTCTTGGTTAAAAATGTTT 292
Db 119 CAGTGGATGACATGACCAAAAGCAAGTATTCATTTTGGTCTCTTGGTTAAAAATGTTT 178

Qy 293 TATTGAAAGTCTTAACCAAAAGCAATATATTTCTGGTGTATGTTTAAATGTTTGTGCAAC 352
Db 179 TGTTCGAAGTCTCAGCACAAGCAATAGCTCTCTAGTATGTTTACTTGGTTCGTCGAGC 238

Qy 353 ATCAATGGGAAAAGCAAGCGCTGGCACTGCCATGTACTAATTTGGAGGAAAGGACTTTA 412
Db 239 ATCAATGGGAAAAGCAAGCGCTGGCACTGCCATGTACTAATTTGGAGGAAAGGACTTTA 298

Qy 413 GTCAAGCTCAAGGAAATGGTGGAGAGCAACTAAATGTTTACTGGAGCAGATGGTTGG 472
Db 299 GTCAAGCTCAAGGAAATGGTGGAGAGCAACTAAATGTTTACTGGAGTATGATGGTTGG 358

Qy 473 TAAAGCCTGTAAATGTGCAACTAACCAAGCTGGAAGAAATTAAGAGAAAT 527
Db 359 TGACTGCTGTAAATGTGCAACTAACCAAGCTGGAAGAAATTAAGAGAAAT 413

RESULT 6
US-10-466-894-375
; Sequence 375, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: US/10/466,894
; PRIOR FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 375
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Homo Sapiens
; NAME/KEY: misc feature
; LOCATION: 422
; OTHER INFORMATION: n = A,T,C or G
; US-10-466-894-375
```


Db 372 TGACTGCTGTAATGTTCAACTAACACAGCTGAAAAA 415
|||||

RESULT 9

US-10-466-894-358
; Sequence 358, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 358
; LENGTH: 486
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 413, 428, 435, 448, 457, 462
; OTHER INFORMATION: n = A,T,C or G

US-10-466-894-358

Query Match 14.1%; Score 284.8; DB 18; Length 486;
Best Local Similarity 89.2%; Pred. No. 2.4e-64;
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 173 CTTTACAAACGAGGCGGAACTACTTTGGACCAAAAGCGAGGACATGGAATGGAAACCA 232
|||||
Db 54 CTTTACAAACGAGGCGGAGACCACTTTGGACCAAAAGCGAGGACATGGAATGGAGAGCG 113
|||||
Qy 233 CAGTGGATGAAATGACCAAAAGCAAGTATTCTTTTGGATTTCTTTGGTTAAAAAATGTT 292
|||||
Db 114 CAGTGGATGACATGACCAAAAGCAAGTATTCTTTTGGATTTCTTTGGTTAAAGAGTGT 173
|||||
Qy 293 TATTGAAAGTCTTAACACAAAGAAATATTTCTGCTGATGTTAATTTGTTTGTGCAAC 352
|||||
Db 174 TGTTGAAAGTCTCAGCACAAAGAAACATAGCTCCTAGTAATGTTACTTGTGTCGAGC 233
|||||
Qy 353 ATGAATGGGAAAGCAAGGCTGGCAGCTGCCATGTACTAAATTTGGAGGAAAGCACTTTA 412
|||||
Db 234 ATGAATGGGAAAGCAAGGCTGGCAGCTGTCTGCTGATTTGGAGGCAAGCACTTTA 293
|||||
Qy 413 GTCAAGCTCAAGGAAATGTTGAGAGGCAACTAAATGTTTACTTGAGCAGATGTTGG 472
|||||
Db 294 GTCAAGCTCAAGGAAATGTTGAGAGGCAACTAAATGTTTACTTGAGTGTGAGTGTGG 353
|||||
Qy 473 TAACAGCTGTATGTGCAACTAACACAGCTGAAAGAAATTA 516
|||||
Db 354 TGACTGCTGTATGTTCAACTAACACAGCTGAAAGAAATTA 397
|||||

RESULT 10

US-10-466-894-428
; Sequence 428, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent

; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 428
; LENGTH: 491
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 12, 421, 441
; OTHER INFORMATION: n = A,T,C or G

US-10-466-894-428

Query Match 14.1%; Score 284.8; DB 18; Length 491;
Best Local Similarity 89.2%; Pred. No. 2.4e-64;
Matches 307; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 173 CTTTACAAACGAGGCGGAACTACTTTGGACCAAAAGCGAGGACATGGAATGGAAACCA 232
|||||
Db 62 CTTTACAAACGAGGCGGAGACCACTTTGGACCAAAAGCGAGGACATGGAATGGAGAGCG 121
|||||
Qy 233 CAGTGGATGAAATGACCAAAAGCAAGTATTCTTTTGGATTTCTTTGGTTAAAAAATGTT 292
|||||
Db 122 CAGTGGATGACATGACCAAAAGCAAGTATTCTTTTGGATTTCTTTGGTTAAAGAGTGT 181
|||||
Qy 293 TATTGAAAGTCTTAACACAAAGAAATATTTCTGCTGATGTTAATTTGTTTGTGCAAC 352
|||||
Db 182 TGTTGAAAGTCTCAGCACAAAGAAACATAGCTCCTAGTAATGTTACTTGTGTCGAGC 241
|||||
Qy 353 ATGAATGGGAAAGCAAGGCTGGCAGCTGCCATGTACTAAATTTGGAGGAAAGCACTTTA 412
|||||
Db 242 ATGAATGGGAAAGCAAGGCTGGCAGCTGTCTGCTGATTTGGAGGCAAGCACTTTA 301
|||||
Qy 413 GTCAAGCTCAAGGAAATGTTGAGAGGCAACTAAATGTTTACTTGAGCAGATGTTGG 472
|||||
Db 302 GTCAAGCTCAAGGAAATGTTGAGAGGCAACTAAATGTTTACTTGAGTGTGAGTGTGG 361
|||||
Qy 473 TAACAGCTGTATGTGCAACTAACACAGCTGAAAGAAATTA 516
|||||
Db 362 TGACTGCTGTATGTTCAACTAACACAGCTGAAAGAAATTA 405
|||||

RESULT 11

US-10-466-894-438
; Sequence 438, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 438

```
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 418
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-438

Query Match      14.1%; Score 284.2; DB 18; Length 420;
Best Local Similarity 87.8%; Pred. No. 3.2e-64;
Matches 310; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 173 CTTTACAAACGAGGCGGAACTACTTTGGACCAAAAGCAGGACATGGAATGGGAAACCA 232
Db      |||
Qy 64 CTTTACAAACGAGGCGGAGCCACCTTTGGACCAAAAGCAGGACATGGAATGGGAGCG 123
Db      |||
Qy 233 CAGTGGATGAATGACCAAAAGCAAGTATTCTTTGGATTCTTTGGTTAAAAAATGTT 292
Db      |||
Qy 124 CAGTGGATGACATGACCAAAAGCAAGTATTCTTTGGATTCTTTGGCTTAAGAAGTGT 183
Db      |||
Qy 293 TATTGGAAGTCTTAACACAAAGCAATATTTCTCTGCTGATGTTAAATGTTGTGCAAC 352
Db      |||
Qy 184 TGTTGGAAGTGTCTCAGCACAAAGCAATAGCTCTTAGTAATGTTACTTTGTTTGTGAGC 243
Db      |||
Qy 353 ATGAATGGGAAAAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGGACTTTA 412
Db      |||
Qy 244 ATGAATGGGAAAAGACCAAGGCTGGCACTGTCTGTCTGCTGATTTGGAGGCAAGGACTTTA 303
Db      |||
Qy 413 GTCAGCTCAAGGAAATGTTGGAGAAAGCACTAAATGTTTACTGGAGCAGATGTTGG 472
Db      |||
Qy 304 GTCACCTCAAGGAAATGTTGGAGAAAGGAGCTAAATGTGTACTGGAGTAGATGTTGG 363
Db      |||
Qy 473 TACAGCTCTTAATGTGCAACTAACACACAGCTGAAGAAATTAACCTAAGAGAA 525
Db      |||
Qy 364 TGACTGCTGTAATGTTCAACTAACACACAGCTGAAGAAATTAACCTAAGAGAA 416
Db      |||

RESULT 12
US-10-466-894-388
; Sequence 388, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 388
; LENGTH: 343
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-466-894-388

Query Match      14.1%; Score 283.8; DB 18; Length 343;
Best Local Similarity 89.2%; Pred. No. 3.6e-64;
Matches 306; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 174 TTTTACACGAGGAGCGGAACTACTTTGGACCAAAAGCAGGACATGGAATGGGAAACCA 233
Db      |||
Qy 1 TTTTACACGAGGAGCGGAGACCACTTTGGACCAAAAGCAGGACATGGAATGGGAGGCGC 60
Db      |||
Qy 234 AGTGGATGAAATGACCAAAAGCAAGTATTCAATTTTGTATTTCTTTGTTAAAAAATGTTT 293
Db      |||
Qy 61 AGTGGATGACATGACCAAAAGCAAGTATTATTTTGTATTTCTTTGTTAAAGAAGTGT 120
Db      |||
Qy 294 ATTTGAAGTCTTAACACAAAGCAATATTTCTCTGCTGATGTTAAATGTTTGTGCAACA 353
Db      |||
Qy 121 GTTTGAAGTGTCTCAGCACAAAGCAATAGCTCTCTAGTAATGTTTACTTTGGTTCTGCGAGCA 180
Db      |||
Qy 354 TGAATGGGAAAAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGGACTTTAG 413
Db      |||
Qy 181 TGAATGGGAAAAGACCAAGGCTGGCACTGTCTGCTGATTTGGAGGCAAGGACTTTAG 240
Db      |||
Qy 414 TCAAGCTCAAGGAAATGTTGGAGAAAGCACTAAATGTTTACTGGAGCAGATGTTGGT 473
Db      |||
Qy 241 TCAACCTCAAGGAAATGTTGGAGAAAGGAGCTAAATGTGTACTGTAGTAGATGTTGGT 300
Db      |||
Qy 474 AACAGCTGTAATGTGCAACTAACACACAGCTGAAGAAATTAAC 516
Db      |||
```

```
Qy 234 AGTGGATGAAATGACCAAAAGCAAGTATTCAATTTTGTATTTCTTTGTTAAAAAATGTTT 293
Db      |||
Qy 61 AGTGGATGACATGACCAAAAGCAAGTATTATTTTGTATTTCTTTGTTAAAGAAGTGT 120
Db      |||
Qy 294 ATTTGAAGTCTTAACACAAAGCAATATTTCTCTGCTGATGTTAAATGTTTGTGCAACA 353
Db      |||
Qy 121 GTTTGAAGTGTCTCAGCACAAAGCAATAGCTCTCTAGTAATGTTTACTTTGGTTCTGCGAGCA 180
Db      |||
Qy 354 TGAATGGGAAAAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGGACTTTAG 413
Db      |||
Qy 181 TGAATGGGAAAAGACCAAGGCTGGCACTGTCTGCTGATTTGGAGGCAAGGACTTTAG 240
Db      |||
Qy 414 TCAAGCTCAAGGAAATGTTGGAGAAAGCACTAAATGTTTACTGGAGCAGATGTTGGT 473
Db      |||
Qy 241 TCAACCTCAAGGAAATGTTGGAGAAAGGAGCTAAATGTGTACTGTAGTAGATGTTGGT 300
Db      |||
Qy 474 AACAGCTGTAATGTGCAACTAACACACAGCTGAAGAAATTAAC 516
Db      |||
Qy 301 GACTGCTGTAATGTTCAACTAACACACAGCTGAAGAAATTAAC 343
Db      |||

RESULT 13
US-10-466-894-390
; Sequence 390, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 390
; LENGTH: 343
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-466-894-390

Query Match      14.1%; Score 283.8; DB 18; Length 343;
Best Local Similarity 89.2%; Pred. No. 3.6e-64;
Matches 306; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 174 TTTTACACGAGGAGCGGAACTACTTTGGACCAAAAGCAGGACATGGAATGGGAAACCA 233
Db      |||
Qy 1 TTTTACACGAGGAGCGGAGACCACTTTGGACCAAAAGCAGGACATGGAATGGGAGGCGC 60
Db      |||
Qy 234 AGTGGATGAAATGACCAAAAGCAAGTATTCAATTTTGTATTTCTTTGTTAAAAAATGTTT 293
Db      |||
Qy 61 AGTGGATGACATGACCAAAAGCAAGTATTATTTTGTATTTCTTTGTTAAAGAAGTGT 120
Db      |||
Qy 294 ATTTGAAGTCTTAACACAAAGCAATATTTCTCTGCTGATGTTAAATGTTTGTGCAACA 353
Db      |||
Qy 121 GTTTGAAGTGTCTCAGCACAAAGCAATAGCTCTCTAGTAATGTTTACTTTGGTTCTGCGAGCA 180
Db      |||
Qy 354 TGAATGGGAAAAGACCAAGGCTGGCACTGCCATGTACTTAATTTGGAGGAAAGGACTTTAG 413
Db      |||
Qy 181 TGAATGGGAAAAGACCAAGGCTGGCACTGTCTGCTGATTTGGAGGCAAGGACTTTAG 240
Db      |||
Qy 414 TCAAGCTCAAGGAAATGTTGGAGAAAGCACTAAATGTTTACTGGAGCAGATGTTGGT 473
Db      |||
Qy 241 TCAACCTCAAGGAAATGTTGGAGAAAGGAGCTAAATGTGTACTGTAGTAGATGTTGGT 300
Db      |||
Qy 474 AACAGCTGTAATGTGCAACTAACACACAGCTGAAGAAATTAAC 516
Db      |||
```

Db 301 GACTGCCTGTAATGTTCACTTAACACAGCTGTAAGAAAAA 343
|||

RESULT 14

US-10-466-894-433
; Sequence 433, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; PRIOR FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 433
; LENGTH: 403
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 218
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-433

Query Match 14.1%; Score 283.8; DB 18; Length 403;
Best Local Similarity 89.0%; Pred. No. 4e-64;
Matches 306; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
Qy 173 CTTTACACGAGGAGCGGAACTACTTTGGACCAAGCGGAGGACATGGAATGGGAACCA 232
Db 56 CTTTACACGAGGAGCGGAGCCACTTTGGACCAAGCGGAGGACATGGAATGGGAGCG 115
Qy 233 CAGTGGATGAATGACCAAAAGCAAGTATTCATTTTGTGATTTCTTTGGTTAAAAATGTT 292
Db 116 CAGTGGATGACATGACCAAAAGCAAGTATTTATTTTGTATTTCTTTGGTTAAGAAGTGT 175
Qy 293 TATTTGAAGTCTTAACACAAAGNATATATTTCTGTGTATGTTAAATGTTTGTGCAAC 352
Db 176 TGTTTGAAGTCTCAGCACAAAGAACATAGCTCCTAGTAATGNTACTTGTGTCGAGC 235
Qy 353 ATGAATGGGAAAAGACCAAGGCTGGCAGCTGCCATGTACTAAATTTGGAGGAAAGACTTTA 412
Db 236 ATGAATGGGAAAGGACCAAGGCTGGCAGCTGTCTGCTGATGTTGGAGGCAAGACTTTA 295
Qy 413 GTCAAGCTCAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGTTGG 472
Db 296 GTCAACCTCAAGGAAATGTTGGAGAGGCAAGTAAATGTTTACTGGAGTAGATGTTGG 355
Qy 473 TAACAGCTGTATGTGCAACTAACACAGCTGAAGAATAAA 516
Db 356 TGACTGCCTGTATGTTCAACTAACACAGCTGAAGAATAAA 399

RESULT 15

US-10-466-894-369
; Sequence 369, Application US/10466894
; Publication No. US20040241671A1
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Tuijnder, Marius
; APPLICANT: Susini, Laurent

; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; FILE REFERENCE: 10918-014-999
; CURRENT APPLICATION NUMBER: US/10/466,894
; PRIOR FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: PCT/FR02/00273
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: FR01/00899
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 2270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 369
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 448,475
; OTHER INFORMATION: n = A,T,C or G
US-10-466-894-369

Query Match 14.0%; Score 283.2; DB 18; Length 497;
Best Local Similarity 89.0%; Pred. No. 6.4e-64;
Matches 306; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
Qy 173 CTTTACACGAGGAGCGGAACTACTTTGGACCAAGCGGAGGACATGGAATGGGAACCA 232
Db 68 CTTTACACGAGGAGCGGAGCCACTTTGGACCAAGCGGAGGACATGGAATGGGAGCG 127
Qy 233 CAGTGGATGAATGACCAAAAGCAAGTATTCATTTTGTGATTTCTTTGGTTAAAAATGTT 292
Db 128 CAGTGGATGACATGACCAAAAGCAAGTATTTATTTTGTATTTCTTTGGTTAAGAAGTGT 187
Qy 293 TATTTGAAGTCTTAACACAAAGNATATATTTCTGTGTATGTTAAATGTTTGTGCAAC 352
Db 188 TGTTTGAAGTCTCAGCACAAAGAACATAGCTCCTAGTAATGTTTGTGTCGAGC 247
Qy 353 ATGAATGGGAAAAGACCAAGGCTGGCAGCTGCCATGTACTAAATTTGGAGGAAAGACTTTA 412
Db 248 ATGAATGGGAAAAGACCAAGGCTGGCAGCTGTCTGCTGATTTGGAGGCAAGACTTTA 307
Qy 413 GTCAAGCTCAAGGAAATGTTGGAGAGGCAACTAAATGTTTACTGGAGCAGATGTTGG 472
Db 308 GTCAACCTCAAGGAAATGTTGGAGAGGCAAGTAAATGTTTACTGGAGTAGATGTTGG 367
Qy 473 TAACAGCTGTATGTGCAACTAACACAGCTGAAGAATAAA 516
Db 368 TGACTGCCTGTATGTTCAACTAACACAGCTGAAGAATAAA 411

Search completed: January 23, 2005, 14:35:20
Job time : 1079 secs

This Page Blank (uspto)